

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_n2p model

Run on: March 22, 2005, 01:51:46 ; Search time 225.596 Seconds
(without alignments)
5517.913 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 3535
Sequence: 1 aggcgcataatagagaag.....ctgaaagggcatagctg99g 1883

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1401741 seqs, 330541175 residues
Total number of hits satisfying chosen parameters: 2803482

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p model -DEV=x1p
-Q=/cgn2/1/USPTO.spool/p/US10030271/runat_21032005.153501.26566/app.query.fasta_1.3150
-DB=Published, Applications AA -QMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdd -LIST=45 -DOCALIGN=200 -THR SCORE=ext -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10030271 @CGN 1.1 438 @runat_21032005.153501.26566
-NCU=6 -ICPU=3 -NO MMAR -LARGQUERY -NEG SCORES=0 -WALT -DSBLOCK=100
-LONGLOG -DEV.TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	1834	51.9	366	14	US-10-106-698-4626	Sequence 4626, Ap
2	1659	46.9	326	14	US-10-296-539-1	Sequence 1, Appl
3	1618	45.8	318	14	US-10-001-254-18	Sequence 18, Appl
4	1605	45.4	319	9	US-09-925-302-758	Sequence 758, App
5	1605	45.4	319	10	US-09-925-302-758	Sequence 758, App
6	1529	43.3	304	14	US-10-013-477-11	Sequence 11, Appl
7	1164	32.9	242	15	US-10-296-115-1220	Sequence 120, Ap
8	999.5	28.3	217	9	US-09-799-777-26	Sequence 26, Appl
9	647.5	18.3	318	9	US-09-935-223-4	Sequence 4, Appl
10	647.5	18.3	318	9	US-09-733-167-1	Sequence 1, Appl
11	641.5	18.1	318	9	US-09-935-223-6	Sequence 6, Appl
12	641.5	18.1	318	9	US-09-733-167-3	Sequence 3, Appl
13	514	14.5	101	14	US-10-001-254-8	Sequence 8, Appl
14	455	12.9	210	9	US-09-733-167-6	Sequence 6, Appl
15	313.5	9.1	19662	15	US-10-084-846A-6	Sequence 6, Appl
16	297.5	8.7	19652	15	US-10-084-846A-7	Sequence 7, Appl
17	290	8.5	19608	15	US-10-084-846A-8	Sequence 8, Appl
18	290	8.5	19695	15	US-10-084-846A-3	Sequence 3, Appl
19	288	8.4	514	16	US-10-437-963-111701	Sequence 11701,
20	278.5	7.9	19608	15	US-10-084-846A-8	Sequence 8, Appl
21	272	7.9	19723	15	US-10-084-846A-5	Sequence 5, Appl
22	268.5	7.8	504	15	US-10-342-331-3	Sequence 3, Appl
23	265.5	7.5	146	14	US-10-001-254-38	Sequence 38, Appl
24	265	7.5	19652	15	US-10-084-846A-7	Sequence 7, Appl
25	265	7.5	19662	15	US-10-084-846A-6	Sequence 6, Appl
26	262.5	7.4	903	14	US-10-156-761-11093	Sequence 11093, A
27	261	7.6	720	15	US-10-342-331-4	Sequence 4, Appl
28	261	7.4	19723	15	US-10-084-846A-5	Sequence 5, Appl
29	259.5	7.3	19725	15	US-10-084-846A-4	Sequence 4, Appl
30	253.5	7.4	1466	15	US-09-918-715-226	Sequence 226, App
31	253.5	7.4	1466	14	US-10-177-293-68	Sequence 68, Appl
32	253.5	7.4	1466	14	US-10-301-822-33	Sequence 33, Appl
33	253.5	7.4	1466	15	US-10-257-021-72	Sequence 72, Appl
34	253.5	7.4	1466	16	US-10-357-851-3	Sequence 3, Appl
35	253.5	7.4	1466	16	US-10-358-024-103	Sequence 103, App
36	253.5	7.4	1466	16	US-10-734-564-103	Sequence 3, Appl
37	250.5	7.1	19656	15	US-10-084-846A-3	Sequence 3, Appl
38	249.5	7.3	1993	17	US-10-488-055-42	Sequence 42, Appl
39	249	7.3	1466	15	US-10-402-089-12	Sequence 12, Appl
40	249	7.0	1466	15	US-10-402-072A-12	Sequence 12, Appl
41	248.5	7.2	422	16	US-10-437-963-136000	Sequence 136000,
42	248	7.0	1953	17	US-10-488-055-42	Sequence 42, Appl
43	247	7.2	19725	15	US-10-084-846A-4	Sequence 4, Appl
44	243.5	7.1	1033	10	US-09-893-519A-14	Sequence 14, Appl
45	238	6.9	625	16	US-10-437-963-165015	Sequence 165015,

ALIGNMENTS

RESULT 1
US-10-106-698-4626
; Sequence 4626, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OR INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: P4005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4626
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-106-698-4626

Alignment Scores:

Pred. No.: 5,37e-96 Length: 366
 Score: 1834.00 Matches: 357
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 51.88% Indels: 0
 DB: 14 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-106-698-4626 (1-366)

QY 28 AAGATTCACAAAACAGTGGCTGGCGCCGTCAGAGATTCAGACGCCACAAATTCGGC 87
 DB 9 LysileGlnAsnLysTrpLeuArgProSerProArgSerHisArgThrProGlnSerGly 28
 QY 88 CGAGGTCTGAGCTGTTCGCCCTCCCTCCCGCGGAATGAGCGGTATCCGAGTTCAGACCCG 147
 DB 29 ArgValLeuSerLeuPheArgLeuProProProGlyMetAlaLeuSerGlySerThrPro 48
 QY 148 GCCCGCTGTGGAGAGAGATGAGTGCCTGAGACTTACCGGATGCTGCTTCACCGT 207
 DB 49 AlaProCytTrpGlnuAspGluCysLeuAspTyrTyrGlyMetLeuSerLeuHisArg 68
 QY 208 ATGTTGAGAGTGTGGCGCGCACTGACCGAGTGCAGCTGAGACTCTCGGCTTCCTG 267
 DB 69 MetPheGlnValAlaGlyGlnLeuThrGluCysGlnLeuGlnLeuAlaPheLeu 88
 QY 268 CTGATGATGAGCTCTCGCGCGCGCGGAGCTTACCGGCGCGAGCGGCTTACAGCTC 327
 DB 89 LeuAspGlnAlaProGlyAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGlnLeu 108
 QY 328 CTGCTGAGACTGAGAGCGCGCGGAGTGCAGCGAGAGAACCTCGGAGTCTGGGCGAA 387
 DB 109 LeuLeuGlnLeuGlnuArgArgGlyGlnCysAspGlnSerAsnLeuArgLeuGln 128
 QY 388 CTCTCGCGGCTGTGCTGGCGCGCGACGACTGCTGGCGCGCGAGCGCGCGCGCG 447
 DB 129 LeuLeuArgValLeuAlaArgHisAspLeuLeuProHisLeuAlaArgLysArgAlaArg 148
 QY 448 CCAGTGTCTCCAGAAAGCTATAGCTATGACACTTCAGCTCTTCAAGAAGAGAGAGGT 507
 DB 149 ProValSerProGlnuArgTyrSerTyrGlyThrSerSerSerLysArgThrGlnGly 168
 QY 508 AACTGCGCTGCGCGCTGCGAGTCAACAGCTTTCGAAATTCACAGAGGCTCAGTGGAG 567
 DB 169 SerCysArgArgArgArgGlnSerSerSerSerAlaAsnSerGlnGlnGlnTrpGln 188
 QY 568 AAGAGCTCCCGCCCAACCAAGCGCGCGCGAGTCCGCGCGCGCGCGAGTGTGTGCC 627
 DB 189 ThrGlySerProProThrLysArgGlnArgArgSerArgGlyArgProSerGlyGlyAla 208
 QY 628 AAGCGCGCGCGAGAGAGGCGCGCGCGAGCGCGAGCGCGAGCGCGAGCGCGAGACT 687
 DB 209 ArgArgArgArgArgArgGlyAlaProAlaAlaProGlnGlnSerGlnuProAlaArgPro 228
 QY 688 TCTCTGAAGGCAAGAGTGAATCTGTGACATCCGCTCCGAGTTCGAGCAGAGTACGCGAG 747
 DB 229 SerSerGlnGlyLysValThrCysAspIleArgLeuArgValArgLaglnuTrpCysGln 248
 QY 748 CATGGCGCGAGCTTGGAGAGAGGCGCTGGATCCCGCGCGCGCGCGCGCGCGAG 807
 DB 249 HisGlyProAlaLeuGlnGlnGlnValAlaSerArgArgProGlnAlaLeuAlaArgGln 268
 QY 808 CTGAGAGCTGTTGGGAGAGCGCAACGAGTGTGCGCTCAAGGGAAGCTGGGCTCTGGGTT 867
 DB 269 LeuAspValAlaPheGlyGlnAlaThrAlaValAlaLeuArgSerArgPhePheLeuValVal 288
 QY 868 TGTGACATCAAGTCTCAGAGCTCTCTATCTGAGAGCGCTTCTGGGCGAGTACCTGAGT 927
 DB 289 CysAspIleLysPheSerGlnLeuSerTyrLeuAspAlaPheTrpGlyLysPyrLeuSer 308
 QY 928 GGGCGCTCTGTGAGGCGCTCGCGGCGCGAGTTCCTGACTGAGGCGCTCGAGAGGCTGTG 987

RESULT 2

US-10-296-539-1

Sequence 1, Application US/10296539

Publication No. US20030165933A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: TANG, Y. Tom

APPLICANT: AZIMZAI, Yalda

APPLICANT: YUE, Henry

APPLICANT: BUREFORD, Neil

APPLICANT: DING, Li

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: PATTERSON, Chandra

APPLICANT: BAUGHN, Mariah R.

TITLE OF INVENTION: REGULATORS OF APOPTOSIS

FILE REFERENCE: PI-0307 PCT

CURRENT APPLICATION NUMBER: US/10/296,539

PRIOR FILING DATE: 2002-11-21

PRIOR APPLICATION NUMBER: 60/209,407; 60/250,326

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PERL Program

SEQ ID NO 1

LENGTH: 326

TYPE: PR

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No. US20030165933A1 3102521CD1

US-10-296-539-1

US-10-030-271-3 (1-1883) x US-10-296-539-1 (1-326)

QY 124 ATGAGCTATCCGAGTTCAGACCGCGCGCGTGTGAGAGAGATGAGTGCCTGAGTAC 183
 DB 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGlnuAspGluCysLeuAspTyr 20
 QY 184 TACGGAGATGCTGTGCTTCCACCGTATGTTCAAGTGTGTGGCGGCGCACTGACGAGTGC 243
 DB 21 TyrGlyMetLeuSerLeuHisArgMetPheGlnValAlaGlyGlnLeuThrGlnCys 40
 QY 244 GAGCTGAGACTCTCTGCGCTTCTGTGCTGATGAGGCTCTCTGGCGCGCGCGAGCTTACGC 303
 DB 41 GlnuGlnLeuLeuAlaPheLeuLeuAspGlnuAlaProGlyAlaAlaGlyGlyLeuAla 60
 QY 304 CGGCGCGCGAGCGCTGAGACTCTGTGAGTGTGAGCGCGCGCGCGAGTGTGCGCGAG 363
 DB 61 ArgAlaArgSerGlyLeuGlnLeuLeuLeuGlnuAspGlnuArgGlyGlnCysAspGln 80
 QY 364 AGCAACTGTGCGCTCTGAGGCGCACTCTGCGCGCTGTGCGCGCGCGCGCAAGCTGTGCGG 423
 DB 81 SerAlaLeuArgLeuLeuGlyGlnLeuLeuArgValAlaAlaArgHisAspLeuLeuPro 100
 QY 424 CACTGGGCGCGCAAGCGCGCGCGCGCAAGTGTCTCCAGAACGCTATGACTATGACCTTCC 483

Db	101	HisLeuAlaIaArgLyAspArgArgProValSerProGluArgTyrSerTyrGlyThrSer	120
QY	484	AGCTTTCAAAGAGACAGAGGGTACCTGCGCTGCGCTGGACATCAAGAGTTCTGCA	543
Db	121	SerSerSerLyAspThrGluGlySerCysArgArgArgGlnSerSerSerAla	140
QY	544	AATTCTCAGCAGGGGTCACTGGGAGACAGCTTCCCTCCCAACCAACGGCAGCGGGAGT	603
Db	141	AsnSerGlnGlnGlyGlnThrIleThrGlySerProProThrLyAspGlnArgArgSer	160
QY	604	CGGGGCGCGCCACAGTGTGTGTGCCAGACGGCGCGCGAGAGGGGGCCCAAGCCGACCCAG	663
Db	161	ArgGlyAspProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaIaProGln	180
QY	664	CAGCAGTCAGAGCCCCGACGACTTCTCTGTAAGGCAAAAGTACCTGTGACATCCGAGCTC	723
Db	181	GlnGlnSerGluProAlaArgProSerSerGluGlyValThrCysAspIleArgLeu	200
QY	724	CGGGTTCAGACAGAGTACTGCGAGCATGGGCCAGCTTGAGACAGGGCGTGGCATCCGG	783
Db	201	ArgValaArgAlaGluTyrCysGluIleGlyProAlaLeuGlnGlnGlyAlaIaSerArg	220
QY	784	CGGGCCCCAGGGGCTGGCGCGGACGTGAGCGTGTGGGCGAGGCCACCGCAGTGTGGCG	843
Db	221	ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg	240
QY	844	TCAAGGAGCACTGGGCTCTGTGTGTGTGATGACATCAAGTTCTCAGAGCTCTCCATCTGAC	903
Db	241	SerArgAspLeuGlySerValValCysAspIleIleYspheSerGluLeuSerTyrLeuAsp	260
QY	904	GCTTCTTGGGCGACTACCTGAGTGTGGCGCCCTGTGTCAGAGCCCTGCGGGCGTGTCTGTG	963
Db	261	AlaPheTrpGlyAspTyrLeuSerSerGlyAlaLeuLeuGlnAlaLeuArgGlyValaPheLeu	280
QY	964	ACTAGAGCGCTCGGAGAGCGTGTGGCGCGGAGAGCGTGTTCGCTGCTGTGATGTGGAT	1022
Db	281	ThrGluAlaLeuArgGluAlaValAlaGlyArgGluAlaValaArgLeuLeuValSerValAsp	300
QY	1024	GAGGCTGACTATGAGAGCTGTGGCGCGCGCGCTGTGTGATGTGAGAGGAAAGAGGGGCGG	1083
Db	301	GluAlaAspTyrGluAlaGlyAlaArgArgArgLeuLeuLeuMetGluGluGlyValArg	320
QY	1084	CGCCCGACAGAGGCGCTCC	1101
Db	321	ArgProThrGluAlaSer	326

RESULT 3

US-10-001-254-18

Sequence 18, Application US/10001254

Publication No. US20030049702A1

GENERAL INFORMATION:

APPLICANT: Reed, John C.

APPLICANT: Godzik, Adam

APPLICANT: Pawlowski, Krzysztof

APPLICANT: Florentino, Loredana

APPLICANT: Lee, Sug Hyung

APPLICANT: Roch, Wilfred

APPLICANT: Stenmer-Liwen, Frank

TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins

FILE REFERENCE: P-LJ 5037

CURRENT APPLICATION NUMBER: US/10/001,254

PRIOR FILING DATE: 2001-11-15

PRIOR APPLICATION NUMBER: 60/301,889

PRIOR FILING DATE: 2001-06-29

PRIOR APPLICATION NUMBER: 09/715,893

NUMBER OF SEQ ID NOS: 62

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 18

LENGTH: 318

TYPE: PRT

ORGANISM: Homo sapien

US-10-001-254-18			
Alignment Scores:			
Pred. No.:	9.52e-84	Length:	318
Score:	1618.00	Matches:	317
Percent Similarity:	99.69%	Conservative:	0
Best Local Similarity:	99.69%	Mismatch:	1
Query Match:	45.77%	Indels:	0
DB:	14	Gaps:	0
US-10-030-271-3 (1-1883) x US-10-001-254-18 (1-318)			
QY	124	ATGGGCGTATCCGGGTGACCCCGGCGCGTGGGAGGAGATGAGTCCCTGGA	183
DB	1	MetalaenusSerglySerThirProLaProCybtrpGluGluAspGlyCysLeuAspTry	20
QY	184	TACGGGATGCTGTGGCTTCAACGTAATGTCAGAGTGGTGGGCGGCAACTGACAGTGC	243
DB	21	TyrGlyMetLeuSerLeuHiAArgMetTheGluValValIGlyGlyGlnLeuThrGluCys	40
QY	244	GAGCTGAGACTCTGGACCTTTCTGCTGGATGAGGCTCTGGCGCGCCGGAGGCTTAGCC	303
DB	41	GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyLeuAla	60
QY	304	CGGGCGCCGACGCGCTTAGACTCTGCTGGAGCTGAGAGCCCGCGGGCAGTGC	363
DB	61	ArgAlaArgSerGlyLeuGluLeuLeuLeuGluLeuGluArgGlyGlyMcysAspGlu	80
QY	364	AGCAACCTGGGGCTGTGGGGCAACTCCTGGCGCGTGGCGCCGCGCAGCACTGGTGGCG	423
DB	81	SerLeuLeuAlaGluLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuPro	100
QY	424	CACCTGGCGCGCAAGCGCGCGCGCGCACTGTCTCCAGAACGCTTAGCTTAGCCACTCC	483
DB	101	HisLeuAlaArgLysArgArgArgProValSerProGluAlaGlyTyrSerTyrGlyThrSer	120
QY	484	AGCTTCTCAANAGAGCAGAGGGTACTCGCGTGGCGCTCGGAGTCAAGCACTTCTGCA	543
DB	121	SerSerSerLysArgThrGluGlySerCysArgArgArgGlnSerSerSerSerAla	140
QY	544	AATTCTCAGCAGGGGTCAAGTGGAGACAGAGTCCCCCGCAACAGCGGACGCGGGAGT	603
DB	141	AsnSerGlnGlnGlyGlnTrpGluThrGlySerProProThrLysArgGlnArgArgSer	160
QY	604	CGGGCGCGCGCCAGTGGTGTGCCAGAGCGGCGGAGAGGGGCCCGCAGCGCACCCGAG	663
DB	161	ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgGlyAlaProAlaAlaProGln	180
QY	664	CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGGCAAAAGTCACTGTGACATCCGAGTC	723
DB	181	GlnGlnSerGluProAlaArgProSerSerGluGlyLysValThrCysAspLysArgLeu	200
QY	724	CGGGTTGACAGACGATCTGGAGCATGGGCCAGCTTGGAGACAGAGGCGTGGCATCCGG	783
DB	201	ArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGluGlnGlyValAlaSerArg	220
QY	784	CGGCGCCAGCGGCTGGCGCGCGCAGCTGACGCGTTTGGGACAGGCGCACCGCAGTGTGGCG	843
DB	221	ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg	240
QY	844	TCAAGGAGCCTGGGCTGTGTGGTTTGTGACATCAAGTTCTGAGAGCTCTCTATCTGAGC	903
DB	241	SerArgAspLeuGlySerValValCysAspLysLeuPheSerGluLeuSerTyrLeuAsp	260
QY	904	GCTTCTGGGGCGACTACCTGATGAGCGCCCTGCTGACAGGCCCTTGGGGGCGTGTCTCG	963
DB	261	AlaPheTrpGlyAspTryLeuSerGlyAlaLeuLeuGlnAlaLeuAlaArgGlyValPheLeu	280
QY	964	ACTAGAGCGCTCGCAGAGGCTGTGGGCGCGGAGGCGTGTTCGCTGCTGATGTGAT	1022
DB	281	ThrGluAlaLeuArgGluAlaValGlyArgGluAlaValArgLeuLeuValSerValAsp	300
QY	1024	GAGCGTGACTTGAAGGCTGTGGCGCGCGCGCTGTGCTGATGATGAGAGGAGG	1077

```

Db      301  GlnAlaAspTyrGlnAlaGlyArgArgLeuLeuMetGlnGlnGly 318
RESULT 4
US-09-925-302-758
; Sequence 758, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIORITY FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIORITY FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIORITY FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-758

Alignment Scores:
Pred. No.:      5,2e-83      Length:      319
Score:          1605.00      Matches:     310
Percent Similarity: 99.36%      Conservative: 0
Best Local Similarity: 99.36%      Mismatches: 1
Query Match:    45.40%      Indels:     1
DB:             9           Gaps:         0

US-10-030-271-3 (1-1883) x US-09-925-302-758 (1-319)
QY      28  AAGATCAAAACAAGTGGTGGCGCGCGCCAGAGATCATCGACGCGCAGAAATCTGAC 87
DB      9  LysIleGlnAsnLysTyrLeuArgProSerProArgSerHisArgThrProGlnSerGly 28
QY      88  CGGGTTCGAGCTGTTCCGCGCTCCCTCCCGCGGAATGAGCGGTATCCGAGTCAACCCG 147
DB      29  ArgValIleuSerLeuPheArgLeuProProGlnMetAlaIleuSerGlySerThrPro 48
QY      148  GCGCCGTGCTGGAGAGAGATGATGCTTCGACTACTACGGAGTGTCTGCCTTACCGT 207
DB      49  AlaProCysTyrGlnGlnAlaAspGlnCysLeuAspTyrTyrGlyMetLeuSerLeuHisArg 68
QY      208  ATGTTGAGGTGGTGGCGGCGCACTGACCGAGTGGAGCTCGTGGCTTTCTG 267
DB      69  MetPheGlnValValGlyGlnLeuThrGlnCysGlnLeuGlnLeuAlaPheLeu 88
QY      268  CTGATGAGAGCTCTGGCGCGCGCGCGAGGCTTAGCCGGCGCGCAGCGCTAGAGCTC 327
DB      89  LeuAspGlnAlaProGlnAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGlnLeu 108
QY      328  CTGCTGGAGCTGAGCGCGCGCGGAGTGGCGGCGAGAGCACTGCGGCTGCTGGGCGAA 387
DB      109  LeuLeuGlnLeuGlnArgArgGlyGlnCysAspGlnSerAsnLeuArgLeuGlnGln 128
QY      388  CTCTCGCGGCGTGGCGCGCGCGCAGCATGCTGCGCGCGCAGCTGGCGGCGCGCGCG 447
DB      129  LeuLeuArgValLeuAlaArgHisAspLeuLeuProHisLeuAlaArgLysArgArgArg 148
QY      448  CCAAGTCTCCAGAAAGCTATAGCTATGACCTTCAGCTTTCAAGAAGAGCAGAGGCT 507
DB      149  ProValSerProGlnArgTyrSerTyrGlyThrSerSerSerLysArgThrGlnGly 168
QY      508  AACTGCGCTGCGCGTGGCGAGTCAAGCAGTTCGAAATTTCTCAGCAGGAGTCAAGGAG 567
DB      169  SerCysArgArgArgArgGlnSerSerSerSerAlaAsnSerGlnGlnGlnTyrGln 188
QY      568  ACAAGCTCCCCCCCAACCAAGCGCGCGCGGAGTGGCGGCGCGCGCGCGCAGTGTGCTGC 627

```

```

Db      169  ThrGlySerProProThrLysArgGlnArgArgSerArgGlyArgProSerGlyGlyAla 208
QY      628  AGACGGCGCGGAGAGGAGGCGCCCGACCCGACCCGACGACGACTAGAGCCCGCAGACT 687
DB      209  ArgArgArgArgArgGlyAlaProAlaAlaProGlnGlnGlnSerGlnProAlaArgPro 228
QY      688  TCCTCTGAAGGCAAGTACCTGTGACATCCGGGCTCCGGGTTTCGAGCAGATGCGCAG 747
DB      229  SerSerGlnGlyLysValThrCysAspIleuArgValArgAlaGlnTyrCysGln 248
QY      748  CATGGGCGAGCCTTGGAGCAGGCGCGTGGCATCCCGCGCGCGCCCAAGCGCTGGCGCGCAG 807
DB      249  HisGlyProAlaLeuGlnGlnGlnGlyValAlaSerArgArgProGlnAlaLeuAlaArgGln 268
QY      808  CTGACAGTGTGTTGGCAGGCCACCCGACGTGCGCTCCTCAGAGGACCTGGGCTCTGTGTT 867
DB      269  LeuAspValAlaPheGlyGlnAlaThrAlaValLeuArgSerArgAspLeuGlySerValVal 288
QY      866  TGTGACATCAAGTTCCTGACAGCTCTCTATCTGACAGCGCTTCTGGGCGGACCTACCTGAGT 927
DB      289  CysAspIleuLysPheSerGlnLeuSerTyrLeuAspAlaPheTyrGlyAspTyrLeuSer 308
QY      928  GCGCGCTGCTGCGAGCGCCCTGGCGGCGGTGTTCC 961
DB      309  GlnAlaLeuLeuGln-ProCysGlyAlaCysSer 319

RESULT 5
US-09-925-302-758
; Sequence 758, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIORITY FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIORITY FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIORITY FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-758

Alignment Scores:
Pred. No.:      5,2e-83      Length:      319
Score:          1605.00      Matches:     310
Percent Similarity: 99.36%      Conservative: 0
Best Local Similarity: 99.36%      Mismatches: 1
Query Match:    45.40%      Indels:     1
DB:             10          Gaps:         0

US-10-030-271-3 (1-1883) x US-09-925-302-758 (1-319)
QY      28  AAGATCAAAACAAGTGGTGGCGCGCGCCAGAGATCATCGACGCGCAGAAATCTGAC 87
DB      9  LysIleGlnAsnLysTyrLeuArgProSerProArgSerHisArgThrProGlnSerGly 28
QY      88  CGGGTTCGAGCTTGTCCGCTCCCTCCCGGGAATGAGCGGTATCCGAGTCAACCCG 147
DB      29  ArgValIleuSerLeuPheArgLeuProProGlnMetAlaIleuSerGlySerThrPro 48
QY      148  GCGCCGTGCTGGAGAGAGATGATGCTTCGACTACTACGGAGTGTCTGCCTTACCGT 207
DB      49  AlaProCysTyrGlnGlnAlaAspGlnCysLeuAspTyrTyrGlyMetLeuSerLeuHisArg 68
QY      208  ATGTTGAGGTGGTGGCGGCGGCACTGACCGAGTGGAGCTCGTGGCTTTCTG 267
DB      69  MetPheGlnValValGlyGlnLeuThrGlnCysGlnLeuGlnLeuAlaPheLeu 88

```



```

Db      181 ---LysSerValThrProaspProLysGluLysGlnThrCysAspIleArgLeuVal 199
QY      730 CGAGCAGACTACTGCGAGCATGGGCGCGCTTGAGACAGAGGCGTGGCATCCGCGGCCC 789
Db      200 ArgAlaGluTyrCysGlnHisGlnThrAlaLeuGlnGlyAsnValPheSerAsnLysGln 219
QY      790 CAGGCGCTGGCGGCGAGCTGAGCTGTTGGGAGGCGCACCGCAGTGTGCGCTCAAG 849
Db      220 AspProLeuGluArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
QY      850 GACCTGGGCTGTGGTTGTGACATGATCTGAGCTCCATCTGACAGCGCTTC 909
Db      240 AspleuGlySerIleIleCysAspIleLysPheSerGluLeuThrTyrLeuAspAlaPhe 259
QY      910 TGGGCGCATCTACCTGAGTGGGCGCGCTGCGACAGCCCTGGGCGCGTGTCTGATGAG 969
Db      260 TPrpArgAspTyrIleAsnGlnSerLeuLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
QY      970 GCCCTGCGAGAGCGTGTGGGCGGAGGCTGTTCGCTGTGATGATGAGCT 1029
Db      280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
QY      1030 GACTATGAGCGTGGCGGCGCGCGCTGTG 1059
Db      300 AspTyrGlnLeuGlyArgGlnLysLeuLeu 309

RESULT 10
US-09-733-167-1
; Sequence 1, Application US/09733167
; Patent No. US2002009909A1
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US2002009909A1 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-733-167-1

Alignment Scores:
Pred. No.: 1,06e-28 Length: 318
Score: 647.50 Matches: 141
Percent Similarity: 63.55% Conservative: 56
Best Local Similarity: 43.48% Mismatches: 92
Query Match: 18.32% Indels: 21
Gaps: 6

US-10-030-271-3 (1-1883) x US-09-733-167-1 (1-318)
QY      157 TGGGAGAGGATAGTGGCTGACCTACTACGAGATGCTGTGCTTACCGCTGATGTTGAG 216
Db      12 TTPPrpGlnGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 31
QY      217 GTGGTGGGGGGGAGCATGACGAGTGGAGCTGAGTGTCTGCGCTTTCG---CTGGAT 273
Db      32 ILeValGlyThrIleuThrIleAspValAlaArgValLeuSerPheLeuPheValAsp 51
QY      274 GAGGCTCTGGCGCGCGGAGGCTTACCGCGGCGCGAGCGCGCTAGAGCTCTGCTG 333
Db      52 ValIleAspAspHisGlnArgLysLeu-----IleArgAsnGlnLysArgAspPheLeu 69
QY      334 GAGCTGAGGCGCGCGGAGCATGTCGGGCGAGAGCAACCTGCGCTGCTGGGCGCACTCTCG 393

```

```

Db      70 AlaLeuGlnArgGlnGlyArgCysAspGlnSerAsnPheArgGlnValLeuLeuLeu 89
QY      394 CCGGTGCTGGCCCGGCACACACTGTGCTGCGACCTGGGCGGAGCGCGGCGGCGAGT 453
Db      90 ArgIleIleThrArgHisAspLeuLeuProTyrValThrLeuLysArgArgAlaVal 109
QY      454 TCTCCA-----GAACTATAGCTATGCGACCTCC-----AGCTCTTCAAG 495
Db      110 CysProAspLeuValAspLysTyrLeuGlnGlnThrSerIleArgTyrValThrProArg 129
QY      496 AGGACAGAGGATAGCTGCGCGCTGCGGAGTCAAGCAGTCTGCAAAATTCAGCAG 555
Db      130 AlaLeuSerAspProGlnProArgProGlnProSerLysThrValProProHisTyr 149
QY      556 GGTACG-----TGGAGACAGGCTCCCGCCCAACCAACGAGCGGCGGAGTGGGCG 609
Db      150 ProValValCysCysProThrSerGlyProGlnMetCysSerLysArgProAlaArgGly 169
QY      610 CCGCCAGTGTGTGTGCCAGACGCGGCGGAGAGGAGGCGCCAGCCGACCCAGCAGCAG 669
Db      170 ArgAlaThrLeuGlySerGlnArgLysArgArg----- 180
QY      670 TGAGAGCCCGCAGACCTTCCTGAGAGGCAAGTGAACCTGTGATCCGCGCTG 729
Db      181 ---LysSerValThrProaspProLysGlnLysGlnThrCysAspIleArgLeuVal 199
QY      730 CGAGCAGACTACTGCGAGCATGGGCGCGCTTGAGACAGAGGCGTGGCATCCGCGGCCC 789
Db      200 ArgAlaGluTyrCysGlnHisGlnThrAlaLeuGlnGlyAsnValPheSerAsnLysGln 219
QY      790 CAGGCGCTGGCGGCGAGCTGAGCTGTTGGGAGGCGCACCGCAGTGTGCGCTCAAG 849
Db      220 AspProLeuGlnArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
QY      850 GACCTGGGCTGTGGTTGTGACATCAAGTCTGAGCTCCATCTGAGCGCTTC 909
Db      240 AspleuGlySerIleIleCysAspIleLysPheSerGlnLeuThrTyrLeuAspAlaPhe 259
QY      910 TGGGCGCATCTACCTGAGTGGGCGCGCTGCGACAGCCCTGGGCGCGTGTCTGATGAG 969
Db      260 TPrpArgAspTyrIleAsnGlnSerLeuLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
QY      970 GCCCTGCGAGAGCGTGTGGGCGGAGGCTGTTCGCTGTGATGATGAGCT 1029
Db      280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
QY      1030 GACTATGAGCGTGGCGGCGCGCGCTGTG 1059
Db      300 AspTyrGlnLeuGlyArgGlnLysLeuLeu 309

RESULT 11
US-09-935-223-6
; Sequence 6, Application US/09935223
; Publication No. US20020086983A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, And
; FILE REFERENCE: TUD2499
; CURRENT APPLICATION NUMBER: US/09/935,223
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 09/723,450
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/276,993
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/859,167
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 318
; TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
US-09-935-223-6

Alignment Scores:
  Pred. No.:      2,32e-28      Length:      318
  Score:          641.50        Matches:      137
  Percent Similarity: 62.90%    Conservative: 58
  Best Local Similarity: 44.19% Mismatches:     94
  Query Match:    18.15%       Indels:       21
  DB:              9           Gaps:         5

US-10-030-271-3 (1-1883) x US-09-935-223-6 (1-318)
QY 157 TGGAGAGAGATGAGTCCCTGACACTAGCATGCTGCTTCACCGTATGTCGAG 216
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 TrpProGluGluArgGlyGluGlnGlnIleGlyLeuTySerLeuHISarGmetPheAsp 31
QY 217 GTGATGGCGCGCAACTGACCGAGTCGAGTCGAGCTGCTGCGCTTCTG---CTGAT 273
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 32 IleValGIyThrHISleuThrHISarGAspValArgValLeuSerPheLeuPheValAsp 51
QY 274 GAGGCTCTGCGCGCGCGAGCTTAGCCCGGCGCGCGAGCGGCTTAGAGCTCTGCTG 333
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 ValIleAspAspHISgluArgGlyLeu-----IleArgAsnGlyArgAspPheLeuLeu 69
QY 334 GAGCTGAGAGCGCGCGCGAGTCGCGGAGAGCAACTGCGCTGCGGCGCAATCTCG 393
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 70 AlLeuGluArgGlnGlyArgCyAspArgIuSerAsnHearGlnValLeuGlnLeuLeu 89
QY 394 CGGTCGTCGCGCGCGCAAGCTGCTGCGCACTGCGCGCAAGCGCGCGCGCGAGT 453
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 90 ArgIleIleThrArgHISAspLeuLeuProTyValThreuleuSylsArgAlaVal 109
QY 454 TCTCCA-----GAACCTATAGCTATGCGACTTCCTCAGCTTTCAAGAGAGAGAG 504
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 CyAspProAspLeuValAspIleSylTyLeuGluGluThrSerIleArgTyValThrProArg 129
QY 505 GGTAGCTCGCGCTGCGCGCGAGTCAGAGCTTCGCAATTCGACAGAGGTGACTG 564
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 AlLeuSerAspProGluProArgProGlnProGlnProSerIleArgTyValThrProHISArg 149
QY 565 GAGACAGAGCTCCCGCGCAACCAAG-----CGGCGCGCGCGAGTGGGCG 609
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 150 ProValValCySerProThrSerGlySerGlnMetCySerIleSylsArgProAlaArgGly 169
QY 610 CGGCGCGAGTGTGTGTCAGACGCGCGCGAGAGAGGGCGCCAGCGCGACCCGAGCAGAG 669
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 ArgThrThreuleuGlySerGlnArgGlyeArgArg----- 180
QY 670 TCAAGCGCGCGCGAGACTTCTCTGAAAGGCAAGTGACTGTGACATCCGGCTCCGGTT 729
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 ---LysSerValThrProAspProTyLeuGluGlnThrCyAspArgIleArgLeuArgVal 199
QY 730 CGAGCAGAGTACTGCGAGCATGGCGCGCGCTTGAGAGAGCGCGGATCCCGCGCGCC 789
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 200 ArgAlaGluIleTyCySerGlnIleGluThrAlaLeuGlnIleAsnValPheSerAsnIleGln 219
QY 790 CAGGCGTCGCGCGCGAGCTGAGCTGTTTGGAGGCGCAACCGAGTGTGCTGCTCAAG 849
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 220 AspProLeuGluArgGlnPheGluArgPheAsnGlnIleAsnThrIleLeuLysSerArg 239
QY 850 GACTGCGCTCTGTGTTTGTGACATCAAGTTCGACAGCTTCCTATTCGAGCGCTTC 909
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 AspLeuGluIleSerIleIleCyAspArgIleLysPheSerIleuThrTyLeuAspAlaPhe 259
QY 910 TGGAGCAGTACTGAGTGGCGCGCTGCTGAGCGCGCTGCGGCGGTTCGTGATGAG 969
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 TrpArgAspTyTrIleAsnGlySerLeuLeuGlnIleuLysGlyValPheIleThrAsp 279
QY 970 GCCCTGCGAGAGGCTGTGGCGCGGAGGCTGTTGCTGCTGCTGAGTGTGATGAGCT 1029
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

Db 280 SerLeuLysGlnAlaValGIyHISgluAlaIleLysLeuLeuValAsnValAspGluGln 299
QY 1030 GACTATGAGGCTGGCGCGCGCGCTGTTG 1059
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 AspTyArgIleuGlnIleArgGlnIleLysLeuLeu 309

RESULT 12
US-09-733-167-3
; Sequence 3, Application US/09733167
; Patent No. US20020099009A1
; GENERAL INFORMATION:
; APPLICANT: Kramer, Peter
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-1120
; CURRENT APPLICATION NUMBER: US/09/733,167
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US20020099009A1 198 25 621.3
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-167-3

Alignment Scores:
  Pred. No.:      2,32e-28      Length:      318
  Score:          641.50        Matches:      137
  Percent Similarity: 62.90%    Conservative: 58
  Best Local Similarity: 44.19% Mismatches:     94
  Query Match:    18.15%       Indels:       21
  DB:              9           Gaps:         5

US-10-030-271-3 (1-1883) x US-09-733-167-3 (1-318)
QY 157 TGGAGAGAGATGAGTCCCTGACACTAGCATGCTGCTTCACCGTATGTCGAG 216
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 12 TrpProGluGluArgGlyGluGlnGlnIleGlyLeuTySerLeuHISarGmetPheAsp 31
QY 217 GTGATGGCGCGCAACTGACCGAGTCGAGTCGAGCTGCTGCGCTTCTG---CTGAT 273
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 32 IleValGIyThrHISleuThrHISarGAspValArgValLeuSerPheLeuPheValAsp 51
QY 274 GAGGCTCTGCGCGCGCGAGCTTAGCCCGGCGCGCGAGCGGCTTAGAGCTCTGCTG 333
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 52 ValIleAspAspHISgluArgGlyLeu-----IleArgAsnGlyArgAspPheLeuLeu 69
QY 334 GAGCTGAGAGCGCGCGCGAGTCGCGGAGAGCAACTGCGCTGCGGCGCAATCTCG 393
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 70 AlLeuGluArgGlnGlyArgCyAspArgIuSerAsnHearGlnValLeuGlnLeuLeu 89
QY 394 CGGTCGTCGCGCGCGCAAGCTGCTGCGCACTGCGCGCAAGCGCGCGCGCGAGT 453
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 90 ArgIleIleThrArgHISAspLeuLeuProTyValThreuleuSylsArgAlaVal 109
QY 454 TCTCCA-----GAACCTATAGCTATGCGACTTCCTCAGCTTTCAAGAGAGAGAG 504
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 CyAspProAspLeuValAspIleSylTyLeuGluGluThrSerIleArgTyValThrProArg 129
QY 110 CyAspProAspLeuValAspIleSylTyLeuGluGluThrSerIleArgTyValThrProArg 129
QY 505 GGTAGCTCGCGCTGCGCGCGAGTCAGAGCTTCGCAATTCGACAGAGGTGACTG 564
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 AlLeuSerAspProGluProArgProGlnProGlnProSerIleArgTyValThrProHISArg 149
QY 565 GAGACAGAGCTCCCGCGCAACCAAG-----CGGCGCGCGCGAGTGGGCG 609
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 150 ProValValCySerProThrSerGlySerGlnMetCySerIleSylsArgProAlaArgGly 169
QY 610 CGGCGCGAGTGTGTGTCAGACGCGCGCGAGAGGGCGCCAGCGCGACCCGAGCAGAG 669
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 170 ArgThrThreuleGlyserGlnArgIysArgArg----- 180
QY 670 TCAGAGCCCGCAGACCTTCTCTGAAAGCAAAGTACCTGTGACATCCGGCTCCGGATT 729
Db 181 ---LysSerValThrProaspProLysGlnIlyGlnThrCysaspIleArgIeuArgVal 199
QY 730 CGAGCAGATCTACTGCGACGATGGGCGCCTTGAGACAGGGCGTGCATCCCGCGGGCC 789
Db 200 ArgAlaGlnIlyrCysGlnIlyrGlnThrAlaIeuGlnIlyrAsnValPheSerAsnIlyrGln 219
QY 790 CAGGCGCTGCGCGCGCTGACGCTGTTGGGCGACCGACCGACGCTGCGCTCAAG 849
Db 220 AspProIeuGlnArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleIeuIlyssrArg 239
QY 850 GACCTGGGCTCTGTGTGTGTGTGACATCAAGTTCAGACCTCTCTATGTGACCGCTTC 909
Db 240 AspIeuGlySerIleIleCysaspIleIlysserSerGlnIleuThrIlyrIeuAspAlaPhe 259
QY 910 TGGGGGACTTACTGTAGTGTGGCGCTGCTGCGAGCGCTTCCGCTGACTGAG 969
Db 260 TTPArgAspIlyrIleAsnGlySerIeuIeuGlnAlaIeuIlyrGlyValPheIleThrAsp 279
QY 970 GCCCTCGAGAGGCTGTGGGCGCGGAGGCTGTTCGCTGTGCTGATGATGAGGCT 1029
Db 280 SerIeuIlyGlnAlaValGlyIlysgIlyAlaIleIlyIeuIeuValAsnValAspGlnIly 299
QY 1030 GACTATGAGCTGTGGCGCGCGCTGTG 1059
Db 300 AspIlyrGlnIleuGlnIlyrGlnIlySerIeuIeu 309

RESULT 13

US-10-001-254-8
; Sequence 8, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sug Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Stenmer-liewen, Frank
; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
; FILE REFERENCE: F-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-254-8

Alignment Scores:

Pred. No.:	3 75e-21	Length:	101
Score:	514.00	Matches:	100
Percent Similarity:	99.01%	Conservative:	0
Best Local Similarity:	99.01%	Mismatches:	1
Query Match:	14.54%	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-001-254-8 (1-101)

QY 157 TGGGAGAGATGATGCTGAGACTACTGACGGAGATCTGTGCTTCAACGATATGTTGAG 216
Db 1 TTPGlnGlnAspGlnIlyrCysIeuAspIlyrIlyrGlyMetIeuSerIeuIlyrAsnIlyrMetPheGln 20
QY 217 GTGGTGGGGGGCACTGACCGAGTGCAGACTGGAGCTCTTGCGCTTTCTGTGATGAG 276

Db 21 ValValIlyGlyGlnIleuThrGlnIlyCysGlnIleuGlnIleuAlaPheIeuIeuAspGln 40
QY 277 GCTCTGGGCGCGCGCGGAGGCTTACCCCGGGCCGACGCGGCTTACGCTCTGTGAG 336
Db 41 AlaProGlyAlaIleGlyGlyIeuAlaArgIlyIleuGlnIleuIeuGln 60
QY 337 CTGAGCGCGCGCGGCACTGCGCGAGACCACTTGGGCTGCTGTGGGCACTCTCTGCG 396
Db 61 IeuGlnArgArgGlyGlnIlyCysAspGlnIlySerAsnIeuArgIleuGlnIlyGlnIleuIlyr 80
QY 397 GTGGTGGCGCGCGCAGACCTGCTGCGGACCTGCGCGGCGGCGGCGGCGGCACTGTCT 456
Db 81 ValIleuAlaArgIlyAspIleuIeuProIlyIleuAlaIlyrGlyArgArgProValSer 100
QY 457 CCA 459
Db 101 Pro 101

RESULT 14

US-09-733-167-6
; Sequence 6, Application US/09733167
; Patent No. US20020099009A1
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US20020099009A1 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C-DEDD, which is a deletion mutant of human DEDD comprising amino
US-09-733-167-6

Alignment Scores:

Pred. No.:	8.61e-18	Length:	210
Score:	455.00 <td>Matches:</td> <td>87</td>	Matches:	87
Percent Similarity:	72.56%	Conservative:	32
Best Local Similarity:	53.05%	Mismatches:	35
Query Match:	12.87%	Indels:	10
DB:	9	Gaps:	1

US-10-030-271-3 (1-1883) x US-09-733-167-6 (1-210)

QY 568 ACAAGCTCCCCCAACCAAGCGGCGGAGCTGGGGCCGCGCCCACTGTGTGCTCC 627
Db 48 ThrSerGlyProGlnMetCysSerIlyrAspProAlaArgIlyrAlaThrIeuIlyser 67
QY 628 AGACGGCGCGGAGAGGGGCGCCACCGACCCAGCGAGCTAGCTAGCGCCCGAGACT 687
Db 68 GlnArgIlyrAspArg-----LysSerValThrPro 77
QY 688 TCCTCTGAAGCAAGTACCTGTGACATCCGCTCCGGGTTTCAGACAGATGTCGAG 747
Db 78 AspProIlyrGlnIlyrGlnThrCysAspIleArgIleuAlaArgAlaGlnIlyrCysGln 97
QY 748 CATGGCAGCCTTGGAGCAGGCGCTGCGCARTCCGCGCGCCCGCGCGCTGCGCGAG 807
Db 98 HisGlnThrAlaIleuGlnIlyrAsnValPheSerAsnIlyrGlnAspProIleuIlyrGln 117
QY 808 CTGAGCGTCTTGGGAGCGGCGCAGCTGCTGCGCTCAAGGAGACTGTGGCTCTGTGTT 867

Dh 118 PheGluArgPheAnGlnAlaAsnThrIleuLysSerArgAspLeuGlySerIleIle 137
Qy 868 TGTCACATCAAGTCTCAGAGCTCTCCATCTGAGCCCTCTGAGGAGTACCTGAGT 927
Dh 138 CysAspIleLysPheSerGluLeuThrTyrIleuAspAlaPheTrpAlaGlyPheIleAsn 157
Qy 928 GGGCCCTGCTGACAGCCCTGCGGGCGGTGTTCTGACTGAGCCCTGCGAGAGGCTGTG 987
Dh 158 GlySerLeuLeuGlnAlaLeuLysGlyValPheIleThrAspSerLeuLysGlnAlaVal 177
Qy 988 GGGCGGAGGCTGTCCCGCTGCTGCTGAGTGTGATGAGCTGACTATGAGGCTGCGCGG 1047
Dh 178 GlyHisGluAlaIleLysLeuLeuValAsnValAspGluGlnAspTyrGluLeuGlyArg 197
Qy 1048 CGCCGCTGTTG 1059
Dh 198 GlnLysLeuLeu 201
RESULT 15
US-10-084-846A-6
; Sequence 6, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MOHLENWEG, AGNES
; APPLICANT: TREFFZER, AXEL
; APPLICANT: BECHTOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 19662
; TYPE: PRN
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-6
Alignment Scores:
Pred. No.: 1,12e-09 Length: 19662
Score: 313.50 Matches: 206
Percent Similarity: 37.04% Conservative: 47
Best Local Similarity: 30.16% Mismatches: 244
Query Match: 9.14% Indels: 189
Dh: 15 Gaps: 36
US-10-030-271-3 (1-1883) x US-10-084-846A-6 (1-19662)
Qy 1698 CCAAGTCAGATCTATGTTGAGCTCTGAGCAGAGGGGGGAAGGTGAGAGTCAAGTTC 1639
Dh 12833 ProGlyProAlaAlaThrAlaProProAspArgArgGlyProAlaArgProSerGlyPro 12852
Qy 1638 TGTAAATGGAAGTGGCAGAGCCCTGCGCTGCTG-----AGATGTGTG 1594
Dh 12853 -----ArgHisGlyArgAlaArgArgGlySerGlyArgArgArgProArgProSer 12869
Qy 1593 GGCAGAGTGAGC-----AGGCCCATGTGACACCCAGCTCAAGTCCCACTGATGT 1543
Dh 12870 GlyArgProAlaAlaProProArgProArgGlyArgProGlySerAlaAlaAspProPro 12889
Qy 1542 GAGCTGAAGCCAGAGTGTGTGTGTCAAATTAGAGGCTCGAGAGAGAGAGCCCTTAGGCC 1483
Dh 12890 Ser---AlaProGly-GlySerSerSerCy6ProGly---GlyArgGlyAlaLeuArgTr 12907
Qy 1482 AAGAAAGTGTGTCTGGGAAGCAGAGCCAGAGCCAGAGGCCAGAGAGATCTCAGCAGAGG 1423

Dh 12907 CARGArgArgProProGlyThrProProAlaArgArgProGlyArgSerProArgArgGly 12927
Qy 1422 CAGGTGAGATAGAGAGAGGTGTGACCTGTGACAGTGCAGAGCAGCCAGATCAGAGA 1263
Dh 12927 Y-----ArgGlyCy6ProGlyThrGlyGlyCy6ArgArgPro----- 12939
Qy 1362 TACAATGTCAGAGGGGCTTTGTGATGAGTGTGA-----GCTGTGCCCTCC 1315
Dh 12940 -AlaCy6Cy6AlaThrAlaAlaGlyArgCy6ProArgArgArgAla-Cy6ProAla 12959
Qy 1314 TTCTGAGATCAGAGCCAGCCCGCTCGGAGGCTTAAGGGGACACCTGGGG----- 1260
Dh 12959 LacGlySerArgArgArgGlyArgAlaArgProLeuAlaProArgArgTrpAsnAlaTr 12979
Qy 1259 --GTAGAGGAGAGTCTGGAGAGAACTCAGCTGGAATGGTTTAGCCTCAGAGCCACA 1202
Dh 12979 hrAlaGlyGlyAlaAlaArgGlyArgAlaAlaGlyLeuArgGlySerArgSerAlaGly 12999
Qy 1201 TCTGTGTGGAGGGCTGTCTCAAGAGGAGTGAAGATGTGCTCTCCAGAGAGAGTGGCCG 1142
Dh 12999 laProThrSerProArgSerArgArgGlySerGlyArgArgProArgArgGlyArg 13019
Qy 1141 GAGACTTGAGGAGTGGATCAATCTGCGAGCTGTGATCAGAGGCTCTGTGCGCGC- 1083
Dh 13019 erGlyProArgAlaGlyCy6AlaVal-----AlaSerValArgArgTr 13033
Qy 1082 --CGCCCCCTTCTCTCTCCATCAGCAGAGGCGCGCCGACAGCTCATGACAGCT 1025
Dh 13033 hrArgProPro-----ArgArgValArgProAlaSerArgGlySer 13047
Qy 1024 CATCC-----ACACTGACAGCAGAGGAGGAGAGCTCCCGC-----CCAGACCT 980
Dh 13047 erAlaArgProArgGlyAlaGlyThrAlaArgArgThr-ProGlyGlyArgProArgPro 13066
Qy 979 CTCGAGAGGCTCAGTCAGAGACACGCGCCGCTGACAGCAGAGGCGCCACTCAGT 920
Dh 13067 -----ProArgSerArgArgAlaAlaProGlyCy6AlaGlySerProProSer 13083
Qy 919 AGTCGCCCAAGAGGCGTCCAGATAGAGAGCTCTGAGAACTTGATGTACAAACCAACAG 860
Dh 13084 ThrArgArgArgArgSerProTrpProAlaGly-----ProArg 13096
Qy 859 AGCCAGAGTCCCTGAGCGCA--GCACCTGCGTGG----- 827
Dh 13097 AlaProSerCy6ArgAlaAlaCy6AlaValArgTrpAspSerAlaArgArgProArg 13116
Qy 826 -----CTGCCCCAAGACGTCCA 809
Dh 13117 AlaAlaGlyArgAlaGlyGlyProCy6AlaGlyCy6GlyArgProArgArgArgPro 13136
Qy 808 GCTGCGCGCGCAGCGCTGTGGGCGCGCGGAGATCCACAGCCCTGCT---CCAGAGCTGACC 752
Dh 13137 AlaValArgProAlaProAlaArgArgSerSerGlyArgArgProSerArgProAlaArgGly 13156
Qy 751 CATGCTGCAGTAC-----TCTGCTCGAACCAG----- 723
Dh 13157 HisArgArgSerArgTrpThrArgSerArgTrpArgTrpProArgProProAlaGlyCy6 13176
Qy 722 -----AGCCGATGTCAAGTCACTTGTGCTTGCAGAGGAAGGTCTGCG- 678
Dh 13177 ProArgAlaAlaArgArgProArgArgArgAlaAspProSerProSerArgHisAlaPro 13196
Qy 677 -----GCTCTGACTGCTGCTGGGTGGGCGGCTGGGCGCTCTCGCGCGCGCTGTG 627
Dh 13197 ArgProArgAlaAlaAlaAlaAlaThrTyrGlyArgSerAlaGlyProArgArgSerArgSer 13216
Qy 626 GCACCAACCACTGGGCGCGCCGACCTCGCGCTGCGCGCTGTGTTGGGGGAGAGCTGTG 567
Dh 13217 AlaPro-----ArgArgArgArgArgArg----- 13224
Qy 566 TCCCACTGACCTGCTGAGAAATTTGCAAACTGCTGACCGGACGCGCAGCTTA 507

```
Db      13225 SerHisArgPro-----SeriGlyArgArgArgAlaArg 13236
QY      506 CCTCTGTCTCTTTGAAGAGCTGAGAGTCATAGCTATAGCTTCTGAGACACTGGC 447
Db      13237 ArgProValThrAlaThrAlaArgAlaArgAlaAlaSerArgSerTrpSeriGly 13256
QY      446 CGGCGCGCGCTTGCGC-----GCCAGGTGGCGGCGAGCTCG----- 411
Db      13257 ArgArgArgProArgCysGlyArgProProAlaAlaSerAlaGlyAlaArgThrArgSer 13276
QY      410 -----TGGCGGCGCGGCGAGCGAGG----- 390
Db      13277 ProProThrProAlaProValaArgArgSerProArgArgArgProAlaArgLeuProAla 13296
QY      389 -----AGTTGCCCGAGAGCGAGCGAGCTCTCGCGCGACCTCGCGCGCGC 342
Db      13297 ValArgGlyProSerArgProThrGlyThrArg-----AspProArgThrProArgArg 13314
QY      341 TCCAGCTCCAGCAGAGCTTAGCGCGCTGCGGCGCGC----- 303
Db      13315 AspGlyArgHisProAlaAlaArgProArgArgAspArgAlaArgValProGluTrpPro 13334
QY      302 -----GCTAAGCGCTCGGCGCGCGAGAGCTCATCCAGCAGAAAGGCCAGAGC 252
Db      13335 AlaArgAlaAlaArgArgProValaArgProGlyArgSer-ProArgArgProProGlyGly 13354
QY      251 TCCA-----GCTGCGACTCGGTCAGTTGCCCGCCGACC 219
Db      13354 yProAlaGlyThrAlaAspGlyProGlyCysAlaArgProArgProGlyGlnArgProPr 13374
QY      218 A-----CCTGAACATACGTTGAAGCAGCAGCATCCCGTAGTAG 180
Db      13374 oArgProAlaArgArgAlaArgProArgGlyHisGlyAlaProProGlyArgProArgTh 13394
QY      179 TCCAGGCGACT-----CATCTCTCCCGAGCAGCGGCGCGGTCGACCCGGAATAGCGCC 126
Db      13394 rProGlyThrGlyPheArgProAlaProAlaArgGly---GlyArgAlaArgTrpAlaPr 13413
QY      125 ATTCCCGGCGGAGG-----AGGCGGAACAAGCTCAGAACCCGGCGAGATTG 79
Db      13413 oValaArgGlyProHisGlyProArgCysArgArgSerArgAsnArgThrGlnProProGly 13433
QY      78 TGGCGTCCGATGACTCTGGGCGAGCGCGCA-----GCCACTGTGTTTGA 32
Db      13433 yGly-GlyAlaAspAlaArgAlaThrAlaGlyProGlySerProAlaGlyCysArgGly 13452
```

Search completed: March 22, 2005, 02:48:05
Job time : 270.596 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2005, 23:44:25 ; Search time 262.015 Seconds
(without alignments)
5558.996 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 3535
Sequence: 1 aggcgcataatagagaag.....ctgaaaggcatagtg995 1883

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues
Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgnt_1/USPTO.spool_P/US10030271/runat_21032005_153457_26448/app_query.fasta_1.3150
-DB=A.GeneSeq_16Dec04 -Qfmt=faetan -SUPFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIG=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pl0 -NDB=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10030271.QCEN_1.1_445 @runat_21032005_153457_26448 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESUBRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.GeneSeq_16Dec04:*
1: geneeqdp1980s:*
2: geneeqdp1990s:*
3: geneeqdp2000s:*
4: geneeqdp2001s:*
5: geneeqdp2002s:*
6: geneeqdp2003as:*
7: geneeqdp2003bs:*
8: geneeqdp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1834	51.9	366	4	AAG73852 Human col
2	1680	47.5	368	5	AB806039 Human NS
3	1666	47.1	326	4	AAB60387 Human apo
4	1659	46.9	326	5	ABB07263 Human apo
5	1659	46.9	326	7	ADC79260 Human DED
6	1618	45.8	318	5	AAE24860 Human DED
7	1618	45.8	318	5	AAE38903 Human DED
8	1605	45.4	319	3	AA858420 Lung canc
9	1574	44.5	361	5	AB806038 Human NS
10	1550	43.8	303	4	AAB60386 Human apo

11	1529	43.3	304	3	AAB15551 Apoptosis
12	1242.5	35.1	277	4	AAAM39805 Human pol
13	1164	32.9	242	4	AAAM25705 Human pro
14	1096.5	31.0	217	4	AAAM1591 Human pro
15	655.5	18.5	318	4	AAAB9404 Human pro
16	647.5	18.3	318	4	AAAB90108 Human pro
17	647.5	18.3	318	5	AAAY51022 Human DED
18	647.5	18.3	318	5	AAE26087 Human FLA
19	646.5	18.3	318	4	AAAB93016 Human pro
20	641.5	18.1	318	2	AAAG90109 Mouse FLA
21	641.5	18.1	318	3	AAAY51023 Murine DE
22	641.5	18.1	318	5	AAE26088 Mouse FLA
23	574	16.2	105	5	ABBP59171 ATP Synth
24	514	14.5	101	5	AAE24855 Human DED
25	514	14.5	101	5	AAE38898 Human DED
26	305.5	8.9	19938	6	ABP76680 Streptomy
27	294	8.6	19938	6	ABP76681 Streptomy
28	286	8.3	1411	7	ABO79662 Streptomy
29	280	8.2	19938	6	ABBB98398 Streptomy
30	273.5	8.0	19938	6	ABP76679 Streptomy
31	268.5	7.8	561	2	AAAB37739 Collagen-
32	268.5	7.8	561	2	AAAB37739 Collagen-
33	268.5	7.8	561	2	AAAB37739 Collagen-
34	264.5	7.7	900	4	ABG03533 Novel hum
35	264	7.7	1040	4	ABG14734 Novel hum
36	262.5	7.7	726	7	ABO71929 Pseudomon
37	261.5	7.6	957	7	ABO71662 Pseudomon
38	261	7.6	777	2	AAAB37740 Collagen-
39	261	7.6	777	2	AAAB37740 Collagen-
40	261	7.6	777	2	AAAB37740 Collagen-
41	260	7.6	681	7	AAAB57651 Collagen-
42	259	7.5	1071	7	ABO75821 Pseudomon
43	257	7.5	450	7	ADP94310 Human col
44	256.5	7.5	450	7	ABO67913 Pseudomon
45	256.5	7.5	19938	6	ABP76678 Streptomy

ALIGNMENTS

RESULT 1	
AAG73852	standard; protein; 366 AA.
AC	AAG73852;
DT	03-SEP-2001 (first entry)
DE	Human colon cancer antigen protein SEQ ID NO:4616.
XX	Human colon cancer; colon cancer antigen; diagnosis; detection;
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX	colorectal carcinoma.
OS	Homo sapiens.
XX	
PN	WO200122920-A2.
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000WO-US026524.
XX	
PR	29-SEP-1999; 99US-0157137P.
XX	
PR	03-NOV-1999; 99US-0163280P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;
XX	
XX	WPI; 2001-235357/24.
DR	N-PSDB; AAH33283.
XX	
XX	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT	useful for preventing, diagnosing and/or treating colorectal cancers.
XX	

PS Claim 11, page 6413-6414; 9803pp: English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon cancer-associated nucleic acid molecules (N) and proteins (P), where the CC proteins are collectively known as colon cancer antigens. The colon CC cancer antigens have cytostatic activity and can be used in gene therapy CC and vaccine production. N and P may be used in the prevention, diagnosis CC and treatment of diseases associated with inappropriate P expression. For CC example, N and P may be used to treat disorders associated with decreased CC expression by rectifying mutations or deletions in a patient's genome CC that affect the activity of P by expressing inactive proteins or to CC supplement the patients own production of P. Additionally, N may be used CC to produce the colon cancer-associated Ps, by inserting the nucleic acid CC into a host cell and culturing the cell to express the proteins. N and P CC can be used in the prevention, diagnosis and treatment of colorectal CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent CC sequences used in the exemplification of the present invention. N.B. CC Pages 666 to 682 and page 7053 of the sequence listing were missing at CC time of publication, meaning no sequences are present for SEQ ID NO:1027 CC to 1052, 7921 and 7922

SQ Sequence 366 AA;

Alignment Scores:

Pred. No.:	3.67e-126	length:	366
Score:	1834.00	Matches:	35
Percent Similarity:	99.72%	Conservative:	0
Best Local Similarity:	99.72%	Mismatches:	1
Query Match:	51.88%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-3 (1-1883) x AAG73852 (1-366)

QY	28	AAAGTCCAAAACAAGTGGCGTGGCGCGCTGGCCACAGAGTCACATGGAGCCCAAGATCTGGC	87
Db	9	Lys1LeGImslnslystrpLeuAargProSerProAArgSerH1sAArgThrpGlnSerCly	28
QY	88	CGGGTTTCGAGCTTGTTCCGCTCCCTCCCTCCCGGGAATGGCGCTATCCGGGTCACACCCG	147
Db	29	ArgValLeuSerLeuPheArgLeuProProProGlnMetAlaLeuSerClySerThrPro	48
QY	148	GCCCCGCTGGAGAGAGATGAGTGCCTTGACCTACACGAGATCTGTGCTTCACCGT	207
Db	49	AlaProCysTrpGlnGlnAerGlnCysLeuAerPylrYrGlnMetLeuSerLeuH1sArg	68
QY	208	ATGTTCGAGGTGGTGGGCGGGCAACTGACCGAGTGCAGCTGGAGCTCTGGCTTTCTG	267
Db	69	MetPheGlnValValGlnGlnGlnLeuThrGlnCysGlnLeuGlnLeuLeuAlaPheLeu	88
QY	268	CTGGATAGAGCTCTGTGGCGCGCGCGAGAGCTTAGCCCGGCGCGGAGGCGCTAGAGCTC	327
Db	89	LeuAspGlnValaProGlnAlaAlaGlnGlnLeuAlaArgAlaArgSerClyLeuGlnLeu	108
QY	328	CTGCTGAGCTGAGCGCGCGCGGAGCTGGCGGAGCAACTGGCGCTGTGGGGCAA	387
Db	109	LeuLeuGlnLeuGlnLuarArgArgGlnCysAspGlnSerAsnLeuAArgLeuGlnGln	128
QY	388	CTCTGGCGCTGCTGGCCCGCCACCACTGGCTGGCGCACTGGCGCGGCAACGGCGCGG	447
Db	129	LeuLeuArgValLeuAlaArgH1sAspLeuLeuProH1sLeuAlaArgLysArgArgArg	148
QY	448	CCAATGCTCTCAGAAAGGCTATGCTATGGCACTCCAGCTCTTCAAAAGAGGAGAGGT	507
Db	149	ProValSerProGlnArgLysTrpSerLysTrpGlnThrSerSerSerLysArgTrpGlnGln	168
QY	508	AGCTGCCGTGGCGCGGCACTCAACAGATTCTGCAAAATTCTCAGACAGAGTCAAGTGGAG	567
Db	169	SerCysArgArgArgArgGlnSerSerSerSerLysAsnSerGlnGlnGlnGlnTrpGln	188
QY	568	ACAGAGTCCCCCAACCAAGCGGACGGCGGAGTGGGAGTGGGCGCCCAAGTGGTGGTCC	627
Db	189	ThrClySerProProThrLysArgGlnAArgSerArgGlnArgProSerGlnGlnVala	208

Oy	628	AGAGGGGGGGGAGAGGGGGCCCGACGGCACGCCAGACGAGAGTGAAGGCCCGACGACT	687
Dd	209	ArgArgArgArgArgArgGGLyAlaProAlaAlaProGlnGlnGlnSerGluProAlaArgPro	228
Oy	688	TCCTCTGAAGAGGCAAGAGTGACTGTGACATCCGGGCTCCGGGTTCGAGCAGATACTGGCAG	747
Dd	229	SerSerGluGluLyIysValThrCysAspIleLeuArgLeuArgValArgAlaGluTyIysGlu	248
Oy	748	CATGGGCGAGCCTTGGAGAGCGGGCGGTGGCATCCGGCGGGCCCGACGGCGCTGGCGGGCAG	807
Dd	249	HisGlyProAlaLeuGluGlnGlnIyValAlaSerArgArgProGlnAlaLeuAlaArgGln	268
Oy	808	CTGAGAGTGTTGGGAGGCGACCGCAGTGGCTCGGCTCAAGGAGCCTGGGCTGTGGTT	867
Dd	269	LeuAspValPheGlyGlnAlaThrAlaValLeuArgSerArgAspLeuGlySerValVal	288
Oy	868	TGTGACATCAAGTTCACAGAGCTCCATCTGACAGCGCTTCTGGGGCGACTACCTGAGT	927
Dd	289	CysAspIleIysPheSerGluLeuSerTyIryLeuAspAlaPheTrpGlyAspTyIryLeuSer	308
Oy	928	GAGCGCCCTGTGTGAGGCGCCCTGGGGGGCGTGTCTGTGATGAGGCCCTCGAGAGGCTGTG	987
Dd	309	GlyAlaLeuLeuGlnAlaLeuArgGlyAlaPheLeuThrGluAlaLeuArgGluAlaVal	328
Oy	988	GGCGGGGAGGCTGTTTGCTGCTGCTGCTGACGTGTGATGAGGCTGACTGAGGCTGGCGGG	1047
Dd	329	GlyArgGluAlaValaArgLeuLeuValSerValaAspGluAlaAspTyIryGluAlaGlyArg	348
Oy	1048	CGCGCGCTGTTGCTGATGTGAGGAGGAGGAGGGGGCGGCGCCGACAGAGGCGCTCC	1107
Dd	349	ArgArgLeuLeuLeuMetGluGluGlnGlyGlyArgArgProThrGluAlaSer	366
RESULT 2			
ABB06039	ID	ABB06039 standard; protein; 368 AA.	
XX	AC	ABE06039;	
XX	DT	10-MAY-2002 (first entry)	
XX	DE	Human NS protein sequence SEQ ID NO:131.	
XX	KM	Human; cytosstatic; osteopathic; gynaecological; neuroprotective;	
KM	KM	antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;	
KM	KM	vasotropic; antiatherosclerotic; antinflammatory; dermatological;	
KM	KM	anorectic; muscular; antiferility; cardiovascular; anticoagulant;	
KM	KM	antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiac;	
KM	KM	anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;	
KM	KM	gastrointestinal; viroicide; antitumor; cerebroprotective; nootropic;	
KM	KM	contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;	
KM	KM	endometriosis; degenerative disease; multiple sclerosis; psoriasis;	
KM	KM	rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;	
KM	KM	inflammation; skin disorder; obesity; muscular dystrophy; AIDS;	
KM	KM	infertility; cardiovascular disease; coagulation disease; hypertension;	
KM	KM	ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;	
KM	KM	diabetes; anxiety; depression; schizophrenia; viral disease; stroke;	
XX	XX	gastric ulcer; Alzheimer's disease.	
OS		Homo sapiens.	
XX	XX		
XX	XX	WO200206315-A2.	
PD	XX	24-JAN-2002.	
PF	XX	17-JUL-2001; 2001MO-II000653.	
XX	XX		
PR	XX	18-JUL-2000; 2000IL-00137345.	
PR	XX	15-DEC-2000; 2000IL-00140354.	
PI	XX	(COMP-) COMPUGEN LTD.	
XX	XX	Mintz L, Freilich S, Bernstein J;	
XX	XX		

PT effector domain and caspase family-cleavage domain, useful in regulating PT diseases with cell proliferation.

PS Claim 3; Page 47-48; 53pp; Japanese.

The invention relates to a novel human apoptosis-associated factor (AAB038d, AAB0387), designated NT2RN1000558, which contains a death effector domain (DED) and a caspase family cleavage domain and is capable of inducing apoptosis in cells. The invention also relates to nucleic acids encoding the protein (AA27407, AA27408); variants of the protein (particularly dominant negative variants); vectors and host cells comprising a nucleic acid which encodes an apoptosis-associated factor of the invention; the recombinant production of the protein; an antibody against the protein; and methods of screening for compounds which can regulate apoptosis. The apoptosis-related factor is useful in regulating diseases associated with cell proliferation and in screening drug candidates e.g., for regulating cell proliferation or cell death in ischemic diseases and chronic viral diseases. The present sequence represents the human apoptosis-associated factor NT2RN1000558

SQ Sequence 326 AA;

Alignment Scores:

Pred. No.:	8.06e-114	Length:	326
Score:	1666.00	Matches:	326
Percent Similarity:	100.00*	Conservative:	0
Best Local Similarity:	100.00*	Mismatches:	0
Query Match:	47.13*	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-3 (1-1883) x AAB60387 (1-326)

QY	124	ATGGGCGCTATCCGGGGTCAGACCCCGGCGCCCGCTGAGGAGAGAGATGATGCTGTGGATAC	183
Db	1	MetAlaIeuSerSerIySerThrProAlaProCysTrpGluGluAspGluCysIleAspTrp	20
QY	184	TACGGAGATGCTGTCCGCTTACCAGTATGTTTGAAGTGTGTGGGGGCAATGACCGAGTGC	243
Db	21	TyrGluMetIeuSerIeuHisArgMetPheGluValValGluGluIleThrThrGluCys	40
QY	244	GACCTGGAGCTCTGTGGCCTTTTCTGCTGGATGAGGCTCTGGGCGCCGGAGGCTTAGCC	303
Db	41	GluLeuGluIleuLeuAlaPheIleuLeuAspGluAlaProGluAlaIleGluIleuAla	60
QY	304	CGGGCCCGCGACGGCTTAAGAGCTCTGCTGAGAGCTGAGACGGCGCGGCGAGCTGGCGGAG	363
Db	61	ArgAlaArgSerCylIeuGluIleuLeuIleuLeuGluIleuAlaArgGluIleCysGlu	80
QY	364	AGCAACCTGTGGCTGTGGGGCAACTCTGCGGTGTGGCGCCGCGACGACTGTGTCGG	423
Db	81	SerAsnIleuArgIeuLeuGluIleuGlnIleuLeuAlaValIleuAlaArgHisAspIleuLeuPro	100
QY	424	CACCTGGCGCGCAAGGGCGGCGCGGCGCAGTGTTCAGAAACGTATAGCTATGGACCTCC	483
Db	101	HisLeuAlaArgIysArgIArgArgProValSerProGluAlaGlyIleSerIyGluIleThrSer	120
QY	484	AGCTCTTCAAAAGAGCAGACAGGGTAGCTCCCTCCGCTCGGCGAGTCAACAGATTTCGA	543
Db	121	SerSerSerIleArgIArgIArgIleuIleSerCysArgArgArgArgGlnSerSerSerIleAla	140
QY	544	AATTCTCAGCAGGGTCAGTGGAGACAGGCTCCCTCCCAACCAACGCGGACGCGGAGT	603
Db	141	AsnSerGlnGlnIleGluIleTrpGluIleThrGlySerProProIleuIleArgGlnAlaArgSer	160
QY	604	CGGGGCGCGGCGCAGGTGTGGTGCAGACGGCGGCGGAGGGGGCCCGACCGCAACCCAG	663
Db	161	ArgGluArgProSerSerIyGluIleAlaArgArgArgArgGluIleAlaProAlaAlaProGln	180
QY	664	CACACAGTCAGAGCCCGCCAGACCTTCTCTTGAAGGCAATGACCTGTGAATCCGGCTC	723
Db	181	GlnGlnSerGluProAlaArgProSerSerGluGluIleValThrCysAspIleArgIleu	200
QY	724	CGGGTTCAGACAGATACCTGCAGACATGGGCAACGCTTGAACAGAGGCGTGGCATCCCGG	783

CC immunological and reproductive disorders. The cell proliferative
 CC disorders include cancers, actinic keratosis, arteriosclerosis,
 CC atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, and
 CC immunological disorders include acquired immunodeficiency syndrome
 CC (AIDS), adult respiratory distress syndrome, Addison's disease,
 CC ankylosing spondylitis, amyloidosis, allergies, anemia, osteoporosis,
 CC autoimmune hemolytic anemia, asthma, autoimmune thyroiditis, Crohn's
 CC disease, contact dermatitis, diabetes mellitus, gout, Graves' disease,
 CC glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
 CC erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis,
 CC urethritis, viral, bacterial, fungal, parasitic, protozoal, helminthic
 CC infections and trauma. Reproductive disorders include disorders of
 CC prolactin production, infertility, endometriosis, polycystic ovary
 CC syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,
 CC disruptions of spermatogenesis, cancer of testis and prostate, impotence,
 CC carcinoma of male breast and gynecomastia. The APRG polynucleotides are
 CC useful for creating knockin humanized animals or transgenic animals to
 CC model human disease and to detect and quantify gene expression in
 CC biopsied tissues in which expression of APRG is correlated with disease.
 CC APRG, fragments of it and antibodies specific for APRG are useful as
 CC elements on a microarray which is useful to monitor or measure protein-
 CC protein interactions, drug-target interactions and gene expression
 CC profiles. The present sequence represents a human APRG polypeptide

XX Sequence 326 AA;

Alignment Scores:

Pred. No.:	2,64e-113	Length:	326
Score:	1659.00	Matches:	325
Percent Similarity:	99.69%	Conservative:	0
Best Local Similarity:	99.69%	Mismatches:	1
Query Match:	46.93%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-3 (1-1883) x ABB07263 (1-326)

QY 124 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 183
 Db 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluuAspGluCysLeuAspTyr 20
 QY 184 TACGGGATGCTGTGCTTACCGGATGTTGAGGTGGGGCGGCACTGACCGAGTGC 243
 Db 21 TylGlyMetLeuSerLeuH1sArgMetPheGluValAlaGlyGlnLeuThrGlnCys 40
 QY 244 GAGCTGGAGGCTCCGGGCTTCTGCTGATGAGGCTCTCGGCGCGCGGAGGCTTACGC 303
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyLeuAla 60
 QY 304 CGGGCCCGCAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCCGCGGAGCTGCGCGAG 363
 Db 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuArgGlyGlnCysAspGlu 80
 QY 364 AGCAACTGCGGCTGCTGGGGCAACTCTGCGCGCTGCGCGCGCACTGCTGCGG 423
 Db 81 SerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgH1sAspLeuLeuPro 100
 QY 424 CACCTGGCGGCGAGCGGCGCGCGGCTGCTGCTGAGAGCTTACTTATGCACTCC 483
 Db 101 H1sLeuAlaArgGlyAspArgArgProValSerProGluArgGlySerTyrGlyThrSer 120
 QY 484 AGCTCTTCAAGAGAGAGAGAGAGTGGCTGCGCGCTGCGCGAGTGAAGCATGTTGCA 543
 Db 121 SerSerSerGlyAspArgThrGlnGlySerCysArgArgArgArgGlnSerSerSerH1s 140
 QY 544 AATTCTCAGCAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGCGCGGAGT 603
 Db 141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 QY 604 CGGGCGCGCGCGCATGAGTGGTGGCAGACGCGGCGGAGAGGCGGCGCGAGCGGCGG 663
 Db 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 664 CAGCAGTCAGAGCGCGCGCAGACTTCTCTGAAAGCAAGTGAAGTGAATCGGCTC 723

Db 181 GlnGlnSerGluProAlaArgProSerSerGluGlyValValThrCysAspH1sLeuLeu 200
 QY 724 CGGGTTCGAGCAGAGTACGCGAGCATGGGCGAGGCTTGGAGAGAGGCGCGATCCCGG 783
 Db 201 ArgValArgAlaGluValCysGlnH1sGlyProAlaLeuGlnGlnGlyAlaAlaSerArg 220
 QY 784 CGGCGCGCGCGCTGCGCGCGCGAGCTGAGCATGTTTGGCGAGGCGCACCGCATGCTGCGC 843
 Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 844 TCAAGGACCTGGGCTGCTGTGTTGTGATCATGATGTTCTCAGAGCTCTCTTATTCGAC 903
 Db 241 SerArgAspLeuGlySerValValCysAspH1sLeuPheSerGluLeuSerTyrLeuAsp 260
 QY 904 GCCTTCGGGCGGACATCACTGAGTGGCGCCCTGCTGACAGGCGCGGCGGCGTTCCTG 963
 Db 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 964 ACTGAGCGCTGCGAGAGGCTGCGCGCGCGGAGGCTTGGCTGCTGATGATGATGAT 1023
 Db 281 ThrGluAlaLeuArgGluAlaValGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 1024 GAGGCTGACTATGAGGCTGCGCGCGCGCTGCTGCTGATGAGAGAGAGAGGCGGCGG 1083
 Db 301 GluAlaAspTyrGluAlaGlyArgArgArgLeuLeuLeuGluGlnGlyGlyArg 320
 QY 1084 CGCGCGCAGAGGCGCTCC 1101
 Db 321 ArgProThrGluAlaSer 326
 RESULT 5
 ID ADC79260 standard; protein; 326 AA.
 AC ADC79260;
 DT 01-JAN-2004 (first entry)
 XX
 DE Human DEDD2 protein SEQ ID NO:2.
 XX
 KW human; death effector domains containing DNA-binding protein;
 KW DED2-containing DNA-binding protein; DEDD2; cell death; gene therapy;
 KW cytoskeletal; cancer; chronic myeloid leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN MO2003054195-A1.
 XX
 PD 03-JUL-2003.
 XX
 PF 20-DEC-2002; 2002MO-JP013371.
 XX
 PR 20-DEC-2001; 2001JP-00387854.
 PR 18-JUL-2002; 2002JP-00209458.
 XX
 PA (MORG) MORINAGA MILK IND CO LTD.
 XX
 PI Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
 DR WPI; 2003-569246/53.
 DR N-PSDB; ADC79259.
 XX
 PT DNA encoding cell death proteins for treatment of kidney, large intestine
 PT and prostate cancers and leukemia.
 PS Claim 1; Page 20-21; 26pp; Japanese.
 XX
 CC The present sequence represents a human death effector domain (DED)
 CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that
 CC causes cell death. Also described: (1) primer and probe for investigation
 CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
 CC cytoskeletal activity. DEDD2 can be used in the diagnosis and treatment of

CC cancers of the kidney, large intestine and prostate, and acute and
 CC chronic myeloid leukaemia.

XX Sequence 326 AA;

SQ

Alignment Scores:
 Pred. No.: 2.64e-113 Length: 326
 Score: 1659.00 Matches: 325
 Percent Similarity: 99.69% Conservative: 0
 Best Local Similarity: 99.69% Mismatches: 1
 Query Match: 46.93% Indels: 0
 DB: 7 Gaps: 0

US-10-030-271-3 (1-1883) x ADC79260 (1-326)

QY 124 ATGGCGCTATCCGGGTGACACCCGCGCGGTGGAGAGAGATGAGTCTGACTAC 183
 DB 1 MetAlaLeuSerIlySerThrProAlaProCySTPGLuInuSpGLuCySLeuAspTyr 20
 QY 184 TACGGATGCTGCTTCAACCGTATGTTGAGTGTGGCGGCACTGACCGAGTGC 243
 DB 21 TYRGLYMetLeuSerLeuHisArgMetPheGLuValAlGlyIyGLuLeuThrGLuCyS 40
 QY 244 GAGCTGAGCTCTGCGCTTTCTTCTGAGTGAAGGCTCTGCGCGCGCGGAGGCTTACGC 303
 DB 41 GLuLeuGLuLeuLeuAlaPheLeuLeuAspGLuAlaProGLyAlaAlaGLyGLyLeuAla 60
 QY 304 CGGCGCCGCGAGCGGCTAGAGCTCTGCTGAGACTGAGAGCGCGCGGCACTGCTGCCAG 363
 DB 61 ArgAlaIArgSerIlyLeuGLuLeuLeuGLuLeuGLuArgIyGLuIyCysAspGLu 80
 QY 364 AGCAACCTGCGGCTGTGGGGCAACTCTCTGCGGCTGTGGCGCGCACTGCTGCCG 423
 DB 81 SerAsnLeuArgLeuLeuGLyGLuLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 424 CACTTGGCGCGAGGCGCGCGCGCGAGTCTCCGAAACGCTATGAGTATGGCACTCC 483
 DB 101 HisLeuAlaArgIyAspArgArgArgProValSerProGLuArgIySerTyrGLyThrSer 120
 QY 484 AGCTTTCAAAGAGACAGAGAGGTAGCTGCGCGCTGCGAGTCAAGCACTTCTGCA 543
 DB 121 SerSerSerIlyArgThrGLuGLuGLySerCyArgArgArgIyGLuLeuSerSerSerAla 140
 QY 544 AATTCTCAGAGGCTGAGTGGAGACAGAGCTCCCGCCCAACCAAGCGGCGCGAGT 603
 DB 141 AsnSerGLuGLuGLuIyGLuIyGLuIyGLuIySerProThrIlyArgIyGLuAlaArgAspSer 160
 QY 604 CGGCGCGCGCGCGAGTGTGGTGCACAGCGCGCGAGAGGGCGCCAGCGCGAG 663
 DB 161 ArgGLyIArgProSerGLyGLyAlaArgArgArgArgIyAlaProAlaAlaProGLu 180
 QY 664 CAGCACTCAGAGCGCGCGAGACTTCTCTGAGAGCAAGTGAAGTGAATCATCCGGCTC 723
 DB 181 GLuGLuSerIyProAlaArgProSerSerGLuGLyIyValThrCySAspIlyLeu 200
 QY 724 CGGTTTGAAGAGAGTACTGCGAGCATGGCGCACTTGGAGAGCGCGTGGCATCCCG 783
 DB 201 ArgValAlaArgAlaGLuIyCysGLuHisGLyProAlaLeuGLuGLuGLuValAlaSerArg 220
 QY 784 CGGCGCGCGCGCGCGCGCGCGAGTGTGGTGTGGAGAGCGCGCGAGTGTGGCG 843
 DB 221 ArgProGLuAlaLeuAlaArgGLuLeuAspValaPheGLyGLuAlaThrAlaValaLeuArg 240
 QY 844 TCAAGGAGACTGGGCTGT 903
 DB 241 SerArgAspLeuIySerValIyCysAspIlyLeuPheSerIyLeuSerIyIyLeuAsp 260
 QY 904 GCCTTGTGGGCGACTACTGAGTGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 963
 DB 261 AlaPheTrpGLyAspIyIyLeuSerGLyAlaLeuLeuGLuAlaLeuArgIyValPheLeu 280
 QY 964 ACTGAGGCGCTGTGAGAGAGGTGTGGGCGCGGAGGCTGTGTGTGTGTGTGTGTGTGT 1023

DB 281 ThrGLuAlaLeuArgGLuAlaValGLyArgGLuAlaValArgLeuLeuValSerValaAsp 300
 QY 1024 GAGCTGACTATGAGCTGTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1083
 DB 301 GLuAlaAspTyrGLuAlaGLyArgArgArgLeuLeuMetGLuGLuGLuIyGLyArg 320
 QY 1084 CGCGCGAGCAGAGCGCTCC 1101
 DB 321 ArgProThrGLuAlaSer 326
 RESULT 6
 ID AAE24860
 XX AAE24860 standard; protein, 318 AA.
 AC AAE24860;
 XX
 XX 22-OCT-2002 (first entry)
 DT
 XX
 XX Human DED4 (death effector domain) protein.
 DE
 XX Human; death domain; DD; death effector domain; DED; Chlamydia infection;
 KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
 KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
 KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
 KW immunosuppressive; gene therapy; antisense therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 53..74
 FT /note="Nuclear localisation sequence"
 XX
 XX WO200240680-A2.
 PN
 XX
 XX 23-MAY-2002.
 PD
 XX
 PF 15-NOV-2001; 2001WO-US044844.
 XX
 XX 17-NOV-2000; 2000US-00715893.
 PR 29-JUN-2001; 2001US-0301889P.
 XX
 XX (BURN-) BURNHAM INST.
 PA
 XX
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Stenmer-Llewellyn F;
 PI
 XX
 DR N-PSDB; AAD40080.
 DR
 XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
 XX
 PS Claim 3; Page 186-187; 209pp; English.
 XX
 XX The invention relates to an isolated polypeptide comprising a death
 CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
 CC is useful for identifying a binding agent, preferably a protein or a drug
 CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
 CC domain from DAP3, IRAK4, CTMD (Chlamydia trachomatis DD protein), DED4 or
 CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
 CC detecting the association of the domain and the candidate binding agent,
 CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
 CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
 CC spectroscopy (MS) and FPA. The invention is useful for modulating the
 CC level of a cell process such as cell proliferation, cell adhesion, cell
 CC stress responses, responses to microbial infection and B cell
 CC immunoglobulin class switching, in particular apoptosis within a cell.
 CC Antibody specifically reactive with CTMD DD of C. trachomatis, C.
 CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
 CC CTMD DD protein is useful for detecting a Chlamydia infection. The
 CC invention is useful for modulating the activity of oncogenic proteins, and
 CC for treating a pathology caused by the oncogenic proteins and for

CC treating bacterial infections by modulating the activity of bacterial
 CC proteins. The protein and antibody specific for it are useful for
 CC discovery of drugs that suppress infection, inflammation, allergy,
 CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
 CC is useful for treating immune-based pathologies, pathologies associated
 CC with cell division, inflammatory diseases such as psoriasis, fibrosis,
 CC arthritis, graft versus host disease. The invention is used in antisense
 CC therapy and gene therapy. The present sequence is human DED4 protein
 XX
 SQ Sequence 318 AA:

Alignment Scores:
 Pred. No.: 2,71e-110 Length: 318
 Score: 1618.00 Matches: 317
 Percent Similarity: 99.69% Conservative: 0
 Best Local Similarity: 99.69% Mismatches: 1
 Query Match: 45.77% Indels: 0
 DB: Gaps: 0

US-10-030-271-3 (1-1883) x AAE24860 (1-318)

QY 124 ATGGCGCTATCCGGGTGAGCCCGCCGCTGGAGAGAGATGAGTGCCTTGACTAC 183
 DB 1 MetAlaLeuSerGlySerThrProAlaProCysTrpIuGluAspGluCysLeuAspTyr 20
 QY 184 TAGCGGATGCTGCTTCAACCCGATGTCGAGGTGGGCGGCGCAACTGACCGAGTGC 243
 DB 21 TrrGlyMetLeuSerLeuHisArgMetPheGluValAlaGlyGlnLeuThrGluCys 40
 QY 244 GAGCTGAGGCTCTGCGCTTCTGCTGATGAGGCTCTCGCGCGCGCGAGGCTTACCC 303
 DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGluLeuAla 60
 QY 304 CGGGCGCGGAGCGGCTTACAGCTCTGCTGAGCTGAGAGCGCGCGGCGAGTGGCGGAG 363
 DB 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuArgGlyGlnCysAspGlu 80
 QY 364 AGCAACCTGGCGGCTGCGGCGCACTCTGCGCGTGGCGCGCGCGCACTGCTGCGG 423
 DB 81 SerHisLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 424 CACCTGGCGCGCAGCGCGCGCGCGCAGTGTCTTCAGAACGCTATAGCACTTCC 483
 DB 101 HisLeuAlaArgLysArgArgArgProAlaSerProGluArgTyrSerTyrGlyThrSer 120
 QY 484 AGCTTTCAAAGAGAGACAGAGGGTAGTCCGCTGCGCGCAGTCAAGCAGTTCTGCA 543
 DB 121 SerSerSerLysArgThrGluGlySerCysArgArgArgArgGlnSerSerSerAla 140
 QY 544 AATTTCAGAGAGGCTCAGTGGGAGACAGGCTCCCGCCCAACCAAGCGGCGCGGAGT 603
 DB 141 AsnSerGlnGlnGlyGlnTrpGluThrGlySerProThrLysArgGlnArgArgSer 160
 QY 604 CGGGCGCGCGCAGTGTGTGTGCAGACGCGCGCGAGAGAGGCGCGCGCGCAGCCAG 663
 DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 664 CAGCAGTCAAGACCCCGCAGACCTTCTTGAAGGCAAGTGAACCTGTGACCTCGGCTC 723
 DB 181 GlnGlnSerGluProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeu 200
 QY 724 CGGGTTGAGAGAGTCTGCGAGCATGGGCGACCTTGGAGCGGGCGGGCGATCCCGG 783
 DB 201 ArgValArgAlaGlyTrpCysGlnHisGlyProAlaLeuGluGlnGlyValAlaSerArg 220
 QY 784 CGGCGCCAGGCGCTGGCGCGCAGCTGAGCTGTTGGGCGAGCGCGCAGTGTCTGCGC 843
 DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 844 TCAGAGGACCTGGGCTCTGAGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGAGC 903
 DB 241 SerArgAspLeuGlySerValValCysAspIleLysPheSerGluLeuSerTyrLeuAsp 260

QY 904 GCCTTGGGCGACATACCTAGTGGCGCCCTGCTGAGCGCCCTGCGGCGGTGTTCTG 963
 DB 261 AlaPheTrpLysAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 964 ACTGAGCGCTTGGAGAGGCTGTGGCGCGGAGGCTGTTGCTGCTGCTGCTGCTGAT 1023
 DB 281 ThrGluAlaLeuArgGluAlaValAlaGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 1024 GAGGCTGACTATAGGCTGGCGCGCGCGCTGCTGCTGATGAGAGGAGGAGG 1077
 DB 301 GluAlaAspTyrGluAlaGlyArgArgArgLeuLeuLeuMetGluGluGluGly 318
 RESULT 7
 AAE38903
 ID AAE38903 standard; protein; 318 AA.
 AC AAE38903;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human DED4 protein.
 XX
 KM Human; death Domain; DD; death effector domain; DED; cell proliferation;
 KM Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;
 KM neural growth factor receptor-interacting death domain; cell adhesion;
 KM vasotropic; microbial infection; inflammation; allograft rejection; CTDD;
 KM cell stress response; benign prostatic hypertrophy; antibacterial; NTDD;
 KM apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;
 KM neoplasia; restenosis; immunosuppressive; antibody therapy; cytotoxic;
 KM keloid.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Domain 12..112
 FT Peptide /note="Death effector domain of DED-4"
 FT 53..74
 FT /note="Nuclear localisation sequence"
 XX
 PN US2003049702-A1.
 XX
 PD 13-MAR-2003.
 XX
 PF 15-NOV-2001; 2001US-00001254.
 XX
 PR 17-NOV-2000; 2000US-00715893.
 PR 17-NOV-2000; 2000US-0367360P.
 PR 29-JUN-2001; 2001US-0301889P.
 XX
 PA (REED/) REED J C.
 PA (GODZ/) GODZIK A.
 PA (PAWL/) PAWLOWSKI K.
 PA (FIOR/) FIORENTINO L.
 PA (LEES/) LEE S H.
 PA (ROTH/) ROTH W.
 PA (STEN/) STENNER-LIEWEN F.
 XX
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Stenner-Liewen F;
 DR WPI; 2002-500222/53.
 DR N-PSDB; AAD59062.
 XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
 XX
 PS Claim 2; Page 44-45; 99pp; English.
 CC The present invention provides novel death Domain (DD) and death effector
 CC domain (DED) proteins and nucleic acids encoding them. The invention also
 CC provides death domain containing protein such as Chlamydia trachomatis
 CC death domain containing protein (CTDD) DD and neural growth factor

receptor-interacting death domain (NIDD). DD. The invention is useful for identifying a binding agent (e.g. protein or drug) that binds a DD, DED or NB-ARC domain from DAB3, IRAK4, CTSD, DED4 or NIDD with a candidate binding agent and identifying an effective agent (e.g. protein or drug) that modulates the association of a DD, DED or NB-ARC domain with protein that binds the DD, DED or NB-ARC domain. The invention is also useful for modulating the level of cell process such as apoptosis, cell adhesion, cell proliferation, cell stress responses, responses to microbial infection and B cell immunoglobulin class switching. DEDs and NB-ARC domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are useful for discovery of drugs that suppress infection, autoimmunity, inflammation, allergy, allograft rejection, sepsis and other diseases. CC inflammation, allergy, allograft rejection, sepsis and other diseases. CC DED or NB-ARC domain proteins are used to treat infection, allergy, autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis, inflammatory hyperplasia and smooth muscle cell proliferation in arteries following balloon angioplasty (restenosis). The invention is also used in antibody therapy and gene therapy. The present sequence is human DED4 protein. The DED4 gene is located on chromosome 19

XX Sequence 318 AA:

Alignment Scores:

Pred. No.:	2	71e-110	Length:	318
Score:	1618.00	Matches:	317	
Percent Similarity:	99.69%	Conservative:	0	
Best Local Similarity:	99.69%	Mismatches:	1	
Query Match:	45.77%	Indels:	0	
DB:	5	Gaps:	0	

US-10-030-271-3 (1-1883) x AAE38903 (1-318)

QY 124 ATGGCGCTATCCGGGTGACACCCCGCGCTGCGAGAGAGATGATGCTGACTAC 183
 Db 1 MetAlaLeuSerGlySerThrProAlaProCysTyrGluGlnuSpGluCysLeuAspTyr 20
 QY 184 TACGGATGTCGTGCTTACCCCTATGTTGAGAGTGTGGGGCGGCACTGACCGAGTGC 243
 Db 21 TygGlyMetLeuSerLeuHisArgMetPheGluValAlaGlyGlyIneuthrGluCys 40
 QY 244 GAGCTGAGCTCTGCGCTTCTTGCTGAGATGAGGCTCTCGGCGCGCGAGGCTTAGCC 303
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyLeuAla 60
 QY 304 CGGCGCCGCGACGCGCTAGAGCTCTGCTGAGACTGAGAGCGCGCGGAGTGGCGAG 363
 Db 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuArgGlyGlnCysAspGlu 80
 QY 364 AGCAACCTGCGGTGCTGGGGCAACTCTCGCGCTGCGCGCGCGCAACCTGCTGCGG 423
 Db 81 SerAsnLeuArgLeuLeuGlnGluLeuLeuValLeuAlaArgHisAspLeuLeuPro 100
 QY 424 CACCTGGCGCGACGCGCGCGCGAGTGTCTCCAGAAAGCATATGATGACCTCC 483
 Db 101 HisLeuAlaArgGlyArgArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
 QY 484 AGCTCTTCAAGAAGACAGAGAGGTAGCTGCGCGCTGCGGAGTCAAGCACTTGTGCA 543
 Db 121 SerSerSerTyrArgThrGluGlySerCysArgArgArgGlnSerSerSerAla 140
 QY 544 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCGCCCAACCAAGGCGGCGCGAGT 603
 Db 141 AsnSerGlnGlnGlyGlnTrpGlnTrpGlySerProThrTyrArgGlnAlaArgAspSer 160
 QY 604 CGGGGCGCGCGCGTGTGGTGCACAGCGCGCGGAGAGGCGCGCGCGCGCGCGCGAG 663
 Db 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 664 CAGCACTCAGAGCCCGCGACAGCTTCTTGAAGAGAAAGTGAAGTGTGACATCCGGCTC 723
 Db 181 GlnGlnSerGlnuProAlaArgProSerSerGlnGlyLysValThrCysAspLeuArgLeu 200
 QY 724 CGGGTTGAGACAGATGCTGCGAGCATGGGCGACGCTTGAGAGCAGGCGCTGCGCGG 783

Db 201 ArgValAlaArgAlaGlyTyrCysGluHisGlyProAlaLeuGlnGlyValAlaSerArg 220
 QY 784 CGGGCCGAGCGCTGCGCGCGAGCTGACGCTGTTGGGCGAGCCACCGCACTGCTGC 843
 Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 844 TCAAGGAGCTGCGCTCTGTGTTGTGACATCAAGTCTCAGAGCTCTCTTATCTGAC 903
 Db 241 SerArgAspLeuGlySerValValCysAspIleTyrPheSerGlnLeuSerTyrLeuAsp 260
 QY 904 GCCTTCTGGGGGAGCTACTGAGTGGCGCGCTGCGAGAGCGCTGCGGGCGGTCTCG 963
 Db 261 AlaPheTyrGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 964 ACTGAGCGCTGCGAGAGCTGCGCGCGGAGGCTGTGCTGCTGCTGCTGCTGAT 1023
 Db 281 ThrGluAlaLeuArgGlnAlaValAlaGlyArgGlnAlaValArgLeuLeuValSerValAsp 300
 QY 1024 GAGGCTGACTATGAGCTGCGCGCGCGCTGTTGATGAGAGGAGAGCGG 1077
 Db 301 GluAlaAspTyrGluAlaGlyArgArgArgLeuLeuMetGluGluGlnGly 318

RESULT 8
 AAB58420
 ID AAB58420 standard; protein; 319 AA.
 XX
 AC AAB58420;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Lung cancer associated polypeptide sequence SEQ ID 758.
 XX
 KW Human; lung cancer associated protein; neuroprotective; cytoskeletal;
 KW cardioactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.
 XX
 OS Homo sapiens.
 XX
 PN WO20005180-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005918.
 XX
 PR 12-MAR-1999; 99US-0124270P.
 XX
 PA (HUMA-) HUMAN GENOME SCT INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Ruben SM;
 XX
 DR WPI; 2000-587514/55.
 DR N-PSDB; AAF18296.
 XX
 PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer.
 XX
 PS Claim 11; Page 1275-1276; 1425pp; English.
 XX
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytoskeletal; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotoxic; antiinfective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the protein
 CC or polynucleotide sequences. The lung cancer associated polynucleotide
 CC sequences may be used for detection of lung cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic

CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, retinosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive

XX Sequence 361 AA:

Alignment Scores:

Pred. No.:	4..8e-107	Length:	361
Score:	1574.00	Matches:	320
Percent Similarity:	96.72%	Conservative:	4
Best Local Similarity:	95.52%	Mismatches:	5
Query Match:	44.53%	Indels:	6
DB:	5	Gaps:	0

US-10-030-271-3 (1-1883) x ABB06038 (1-361)

QY 85 GGGCGGATTCAGAGCTTTGCGGCTCCCGCGGAAATGGCGCTATCCGGGTCGACC 144
 DB 16 GtAArgValLeuSerLeuPheArgLeuProProGlyMetAlaLeuSerGlySerThr 35
 QY 145 CCGGCCCCGTCGGAGAGGATGAGTGCCTGAGTACTACTACCGGATGCTGTGGCTTCAC 204
 DB 36 ProAlaProCysTrpTrpGluAspGluCysLeuAspTrpTrpGlyMetLeuSerLeuHis 55
 QY 205 CGTATGTCAGAGTGGTGGCGCGCAACTGACCGAGTGCAGAGTGCAGTCTCGGCTTT 264
 DB 56 ArgMetTrpGluValValGlyGlyGlnLeuThrGluCysLeuSerLeuLeuAlaPhe 75
 QY 265 CTGCTGATGAGGCTCTCGCGCGCGCGGCTTACCGCGGCGCGGCGCTTACGAG 324
 DB 76 LeuLeuAspGluAlaProGlyAlaAlaGlyGlyTyrAlaArgAlaArgSerGlyLeuGlu 95
 QY 325 CTCCTCTGAGAGTGCAGCGCCCGCGGCACTGGCGGAGGCACTCGCGGCTGCTGGG 384
 DB 96 LeuLeuLeuGlnLeuGlnArgGlnGlnGlnCysAspGlnSerAsnLeuArgLeuGln 115
 QY 385 CAACCTCGCGCGCTGCGCGCGCGCAAGCTCTGCGCGCACTGGCGCGCGCGCGC 444
 DB 116 GlnLeuLeuArgValAlaLeuAlaGlnHisAspLeuLeuProHisLeuAlaArgLys-AlaAla 135
 QY 445 CGGCGAGTGTCTCCAGAACGCTTATGAGTACCTGACCTCTTCAAGAGACAGAG 504
 DB 135 aglyGlnCysLeuGlnHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 155
 QY 505 GGTAGTGCCTGCGCGCGCGCGCAAGTTCGCAATTCGACGAGGCTCAGTGG 564
 DB 155 gValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 175
 QY 565 GAGACAGAGCTCCCGCCCAACCAAGCGGAGCGGAGTGGGGCGCGCCAGTGGTGT 624
 DB 175 yArgGlnAlaProProGlnProSerGlyThr-AlaGlnSerGlyProAlaGlnTrpTrpC 195
 QY 625 GCCAGACGCGCGCGGAGAGGCGCGCGCAAGCGGCAAGGAGAGTCAAGCGCGCA 684
 DB 195 ySerGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 214
 QY 685 CCTTCTCTGAGAGCAAGTACCTGAGATCCGGCTCGGGTTCGAGAGTACTGC 744
 DB 215 ProSerSerGlnGlyValThrCysAspAlaArgLeuArgValAlaAlaAlaAlaAla 234
 QY 745 GAGCATGGGCGAGCTTGGAGAGGCGGTGAGTCCCGGCGCGCGCGCGCGCTGGCGCG 804
 DB 235 GlnHisGlyProAlaLeuGlnGlnGlyValAlaSerArgArgProGlnAlaAlaAlaAla 254
 QY 805 CAGCTGAGCGTGTGGGAGGCGCACCGAGTCTGGCTCAAGGAGCCTGGGCTGTGTG 864
 DB 255 GlnLeuAspValAlaPheGlyGlnAlaTrpAlaValLeuArgSerArgAspLeuGlnGlySerVal 274
 QY 865 GTTTGAGCATCAAGTTCTCAGAGCTCTCTATCTGGAGCGCTTGGGGCGACTACCTG 924

DB 275 ValCysAspIleuLeuPheSerGlnLeuSerTrpLeuAspAlaPheTrpGlyAspTrpLeu 294
 QY 925 AGTGGCGCCCTGTCGAGGCGCTGGCGGCGGCTGTTCTGACTAGAGCCCTGGAGAGGCT 984
 DB 295 SerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeuThrGluAlaLeuArgGluAla 314
 QY 985 GTGGGCGCGGAGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
 DB 315 ValGlyArgGlu-LeuPheAlaCysTrpSerValTrpMetArgLeuThrMetArgLeuAl 334
 QY 1045 CGGCGCGCGCTGTTGCTGAGAGGAGAGGAG 1077
 DB 334 a-AlaProProValAlaAsp-GluGlnGlyGly 344
 RESULT 10
 AAB60386
 ID AAB60386 standard; protein; 303 AA.
 AC AAB60386;
 AC 24-APR-2001 (first entry)
 DT 24-APR-2001 (first entry)
 DT 24-APR-2001 (first entry)
 DE Human apoptosis-associated factor NT2RM1000558 (portion), SEQ ID NO:2.
 DE Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
 KW DED; caspase family cleavage domain; pro-apoptotic; drug screening;
 KW cell proliferation; ischaemic disease; chronic viral disease.
 OS Homo sapiens.
 OS Homo sapiens.
 OS Homo sapiens.
 PN MO200104300-A1.
 PD 18-JAN-2001.
 PD 18-JAN-2001.
 PF 06-JUL-2000; 2000MO-JP004516.
 PR 08-JUL-1999; 99JP-00194179.
 PR 18-OCT-1999; 99US-0159586P.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
 PI MPI; 2001-138348/14.
 DR N-PSDB; AAF27407.
 DR N-PSDB; AAF27407.
 XX Polynucleotide encoding an apoptosis-associated factor protein with death
 PT effector domain and caspase family-cleavage domain, useful in regulating
 PT diseases with cell proliferation.
 PS Claim 1; Page 43-44; 53pp; Japanese.
 XX The invention relates to a novel human apoptosis-associated factor
 CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
 CC effector domain (DED), and a caspase family cleavage domain and is capable
 CC of inducing apoptosis in cells. The invention also relates to nucleic
 CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
 CC (particularly dominant negative variants); vectors and host cells
 CC comprising a nucleic acid which encodes an apoptosis- associated factor
 CC of the invention; the recombinant production of the protein; an antibody
 CC against the protein; and methods of screening for compounds which can
 CC regulate apoptosis. The apoptosis-related factor is useful in regulating
 CC diseases associated with cell proliferation and in screening drug
 CC candidates e.g., for regulating cell proliferation or cell death in
 CC ischemic diseases and chronic viral diseases. The present sequence
 CC represents a substantial proportion of the human apoptosis-associated
 CC factor NT2RM1000558
 XX
 XX Sequence 303 AA;
 Alignment Scores:
 Pred. No.: 2.68e-105 Length: 303

Score: 1550.00 Matches: 303
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.85% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x AAB60386 (1-303)

QY 124 ATGGCGCTATCCGGGCTGACCCCGGCGCCGCTGCTGGAGAGAGATGAGTCCGACTAC 183
Db 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluGlnAaspGluCysLeuAAspTyr 20
QY 184 TAGGGATGCTGTGCTTACCCGATGTTGAGGTTGGCGGGGCACTGACCGAGTGC 243
Db 21 TyrGlyMetLeuSerLeuH1sArgMetPheGluValAlaGlyGlyGlnLeuThrGluCys 40
QY 244 GAGCTGAGGCTCCGCTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTACGC 303
Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60
QY 304 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGGCGGAG 363
Db 61 ArgAlaArgSerGlyLeuGluLeuLeuGlnLeuGlnLeuArgGlyGlnCysGlyGlu 80
QY 364 AGCAACTCGGCTGCTGGGCGCACTCTGCGCGCTGCGCGCGCGCACTGCTGCGG 423
Db 81 SerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgH1sAAspLeuPro 100
QY 424 CACCTGGCGGCGAAGCGCGCGCGGCTCTCCAGAGCTTACCTAGTGGACCTCC 483
Db 101 HisLeuAlaArgGlySarGlyArgArgProAlaSerProGluArgTyrSerTyrGlyThrSer 120
QY 484 AGCTCTTCAAG 543
Db 121 SerSerSerTyrSerThrGlnGlySerCysArgArgArgArgGlnSerSerSerAla 140
QY 544 AATTCTCAGAGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 603
Db 141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
QY 604 CGGGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
Db 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgArgArgArgArgArgArg 180
QY 664 CAGCAGTCAAG 723
Db 181 GlnGlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGlnGln 200
QY 724 CGGGTTCAG 783
Db 201 ArgValArgAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
QY 784 CGGCCCCAGGCGGCTGCGCGGCGGAGCTGAGCTGTTGGGAGAGAGAGAGAGAG 843
Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAAspValPheGlyGlnAlaThrAlaValLeuArg 240
QY 844 TCAAGGAGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
Db 241 SerArgAspLeuGlySerValValCysAAspIleLeuSerGlnLeuSerTyrLeuAAsp 260
QY 904 GCCTTCTGAGGAG 963
Db 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
QY 964 ACTGAGGAG 1023
Db 281 ThrGlnAlaLeuArgGlnAlaValAlaGlyArgGlnAlaValArgLeuLeuValSerValAsp 300
QY 1024 GAGGCTGAGC 1032
Db 301 GluAlaAsp 303
RESULT 11

AAB15551
ID AAB15551 standard; protein; 304 AA.

AC AAB15551;

DT 28-FEB-2001 (first entry)

DE Apoptosis related protein encoded by gene 1 clone H1DOK36.

KW Cytostatic; antiinflammatory; immunosuppressive; antisclerotic; cardiant;
KW viral; anti-AIDS; vasotropic; anti-ischaemic; antiparkinsonian;
KW anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;
KW colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;
KW autoimmune disorder; multiple sclerosis; viral infection.

OS Homo sapiens.

PN NO200056752-A2.

PD 28-SEP-2000.

PF 15-MAR-2000; 2000MO-US006642.

PR 24-MAR-1999; 99US-0126018P.

PR 17-JUN-1999; 99US-0139638P.

PR 18-AUG-1999; 99US-0149449P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Young PA;

XX WPI; 2000-587660/55.

DR N-PSDB; AAA95790.

PT Nucleic acids encoding human apoptosis associated protein, useful for the

PT prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's

PT disease, inflammation and ischemic injury.

PS Claim 11, Page 252-253; 273pp; English.

XX The invention relates to the isolation of genes encoding 9 human

CC apoptosis-related proteins. The nucleotide sequences AAA95790-A95798

CC encode the human apoptosis related proteins AAB15551-B15559. The genes

CC can be used to generate fusion proteins by linking to the gene for the

CC human immunoglobulin G Fc (IgG Fc) portion (AAA95799) for increasing the

CC stability of the fusion protein as compared to the human protein only.

CC The gene and encoded protein may be used in the prevention, treatment and

CC diagnosis of diseases associated with inappropriate apoptosis associated

CC protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,

CC melanomas and lymphomas), inflammation, autoimmune disorders (e.g.

CC multiple sclerosis) and viral infections (e.g. herpes)

XX Sequence 304 AA;

SQ

Alignment Scores:

Pred. No.: 9.39e-104 Length: 304
Score: 1529.00 Matches: 303
Percent Similarity: 99.67% Conservative: 0
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 43.25% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x AAB15551 (1-304)

QY 190 ATGCTGTGCTTACCGGTATGTTGAGGTGGGCGGCACTGACCGAGTGGAGCTG 249

Db 1 MetLeuSerLeuH1sArgMetPheGluValAlaGlyGlyGlnLeuThrGluCysGluLeu 20

QY 250 GAGCTCTGCGCTTCTGCTGATGAGGCTCTGCGGCGCGGAGAGCTTACGCGGCGC 309

Db 21 GluLeuLeuAlaPheLeuLeuAAspGlnAlaProGlyAlaAlaGlyGlyLeuAlaArgAla 40

QY 310 CGCAGCGGCTAGAGCTCTGCTGAGCTGAGAGCGCGGCGGAGTGGCGAGAGCAAC 369


```

Db      61 ArgAlaArgSerGlyLeuLysLeuLeuLeuLLeuLLeuLuhArgArgGlyGlnCyAspGlu 80
Qy      364 AGCAACTGGGCGGTGCTGGGGCAACTCCTGGCGGTGGCCGCGCAAGACTCTGCTCGG 423
Db      81 SerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
Qy      424 CACCTGGCGCGCAAGCGCGCGCGCGAGTGTCTTCAGAACGCTATAGCTATGGCACTCC 483
Db      101 HisLeuAlaArgLysArgArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
Qy      484 AGCTCTTCAAGAGAGACAGAGAGGTAGTCCGCTGCGGTGGAGCTCAAGCATGTTCTGCA 543
Db      121 SerSerSerLysArgThrGlnGlySerCyAspArgArgArgGlnSerSerSerAla 140
Qy      544 AATTCTCAGCAGGGGTCACTGGGAGACAGAGCTCCCCCAACCAAGCGGCAAGCGGAGT 603
Db      141 AsnSerGlnGln-----GlySerProProThrLysAspGlnAspArgSer 155
Qy      604 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCCGACCGCACTCCAG 663
Db      156 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 175
Qy      664 CAGCAGTCAGAGCCCGCGCAAGCTTCTCTGGAAGCAAGTGACCTGTGATCCGGCTC 723
Db      176 GlnGlnSerGluProAlaArgProSerSerGlnGlyLys----- 188
Qy      724 CGGGTTGAGACAGAGTACTGGAGCATGGGCGAGCTTGGAGCAGGGCGTGGCATCCCG 783
Db      188 ----- 188
Qy      784 CGGCGCCAGCGCTGGCGCGCGAGCTGAGCGTGTGGGCGACCGACCGAGTGTGCGC 843
Db      188 ----- 188
Qy      844 TCAAGGACCTGGGCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 903
Db      188 ----- 188
Qy      904 GCCTTCTGGGGCGACTACTGAGTGGGCGCCCTGTGAGAGCCCTGGGGGGGTGTCTCG 963
Db      188 ----- 188
Qy      964 ACTGAGCCCTGGCGAGAGGCTGTGGGCGGGGAGGCTGTGCTGTGATGATGTGAT 1023
Db      188 ----- 188
Qy      1024 GAGGCTGACTATGAGCTGGCGCGCGCTGTGCTGATGAGAGAGAAAGGGGGCGG 1083
Db      188 ----- 188
Qy      1084 CGCCCGACAGAGGCTCTGTATCCAGGACTGGCAGATGTATCCCACTCCAAGTCTCG 1143
Db      188 ----- 188
Qy      1144 GGCACACTTCTCTGGAGAGAGCAGCATCTTACCTCTGACAGCCCTCCACAGGATG 1203
Db      188 ----- 188
Qy      1204 TGGGCTCTGAGGCTTAACATTTCCAGCTGATGTTCTTCCAGACTCCTCTACCCCC 1263
Db      188 ----- 188
Qy      1264 AGGTGTGCCCCCTTAGCTCCGAGGCGGGGGCTGGGCTGTATCTCAGAAAGGAGGGGC 1323
Db      188 ----- 188
Qy      1324 ACAGCTACAGACTCACCAAGGCCCGCTGACATTTGATCTGTGATCTTGGGCTGTCTG 1383
Db      188 ----- 188
Qy      1384 CACTGTACAGGTGACACACTGCTCATGCTACACTGCCCTGCTGAGATCTTCCCTG 1443
Db      188 ----- 188

```

```

Qy      1444 GGCCTTGCCCTGGGCTGCTTCCAGACACACTTCTTGGCTTAAGGCTTCTCTCA 1503
Db      188 ----- 188
Qy      1504 GGAAGCTTAATTTGACCACCAACCAAGCTGGGCTTACGACCATCATGAGGCACTGAGCT 1563
Db      189 -----Val--ThrThrThrAsnLeuGlyPheSerHisIleSerGlyHisIleTrpSer 205
Qy      1564 GGGGTGACATTTGGGGCTGTGCTCACTTGGCCCAACATCTCCAGCCAGCGGCGCTGGCC 1623
Db      205 rpglyAlaHisGlyAlaCySerSerProCySerProHisIleSerSerGlnProGlyProCyAsp 225
Qy      1624 CAGCTTCAATTTTACAGACCTGACTCTCTCAGCTTCCCCCTGCTGTCAGACTGACA 1683
Db      225 roAlaSerIleTyrArgProAspSerProHisLeuProProCySerProGluLeuAsnI 245
-Qy      1684 TAGACTTGACCTTGATGTGACCTGAGTGCATGAGAGTGTATGAGAGCATCATAC 1743
Db      245 leAspLeuHisLeuAspValThrTrpSerValThrTrpGluCySerGlySerIleIleP 265
Qy      1744 CAAGG 1748
Db      265 roArg 266

RESULT 13
ID      AAM25705
AC      AAM25705;
DT      16-OCT-2001 (first entry)
DE      Human protein sequence SEQ ID NO:1220.
OS      Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW      antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW      antibacterial; endocrine; cardiact; central nervous system; virucide;
KW      anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;
KW      antiaggregant; haemostatic; vulnery; antidiabetic; eczema;
KW      dermatological; antiallergic; antistimulant; antidiabetic; cyrostatic;
KW      neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW      immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;
KW      antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW      cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW      genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW      thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW      allergic rhinitis; diabetes; multiple sclerosis; depression;
KW      Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW      neurological disorder.
OS      Homo sapiens.
PN      WO200153455-A2.
PD      26-JUL-2001.
PF      22-DEC-2000; 2000MO-US035017.
PR      23-DEC-1999; 99US-00471275.
PR      21-JAN-2000; 2000US-00488725.
PR      25-APR-2000; 2000US-00552317.
PA      (HYSE-) HYSEQ INC.
XX      Tang YT, Liu C, Drmanac RT;
PI      WPI; 2001-457603/49.
XX      DR      N-PSDB; AAH99646.
XX      Isolated human polynucleotides encoding polypeptides, useful for the
PT      treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX

```

PS Claim 20; Page 253; 1217bp; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antineoplastic; antineoplastic;
 CC antineoplastic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; virucide; anti-HIV; fungicide; antitumor;
 CC cardiovascular; antineoplastic; antineoplastic; antineoplastic;
 CC antineoplastic; osteoporosis; dermatological; antineoplastic;
 CC antineoplastic; cytotoxic; neuroprotective; antidepressant; nootropic;
 CC antiparasitic; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders

XX Sequence 242 AA;

Alignment Scores:

Pred. No.: 6.14e-77 Length: 242
 Score: 1164.00 Matches: 241
 Percent Similarity: 66.48% Conservativity: 1
 Best Local Similarity: 66.21% Mismatches: 0
 Query Match: 32.93% Indels: 122
 DB: 4 Gaps: 1

US-10-030-271-3 (1-1883) x AAM25705 (1-242)

QY 10 AATAGAGAGCTGCCGAAAGATCCAAACAGTGGCTGCGCCGCCAGAGTCAAT 69
 DB 1 AATAGAGAGCTGCCGAAAGATCCAAACAGTGGCTGCGCCGCCAGAGTCAAT 20
 QY 70 CGGACCGCAAGTCTGGCGGGTTTGAGCTTCTCCCTCCCTCCCGGGAATGGCG 129
 DB 21 AATAGAGAGCTGCCGAAAGATCCAAACAGTGGCTGCGCCGCCAGAGTCAAT 25
 QY 130 CTATCGGGTTCAGACCCCGCCCGTGTGGAGAGAGATGAGTCTGGAATAGAGG 189
 DB 25 ----- 25
 QY 190 ATGCTGCTGCTTACCCGTATGTTGAGGTGTGGGGGCACTGACCAAGTCCAGCTG 249
 DB 25 ----- 25
 QY 250 GAGCTCTGAGCTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAGCCCGGGCC 309
 DB 25 ----- 25
 QY 310 CGGACCGCTAGAGCTCTGCTGAGCTGAGAGCCCGCGGCAATGCGCGAGAGCAAC 369
 DB 25 ----- 25
 QY 370 CTGCGGCTGTGGGGCACTCTGCGCGGTGTGGCCCGCAGACCTGCTGCGCACTG 429
 DB 25 ----- 25
 QY 430 GCGCGAAGCGGCGCGGCAAGTCTCCAGAAAGCTATAGTATGAGCACTTCACTCT 489
 DB 26 -----ValSerProGluArgTyrSerTyrAlaThrSerSerSer 38
 QY 490 TCAAGAGAGCAAGAGGAGTGTGCGCGCGGTGCGAGTCAAGAGTCTGCAAAATTC 549
 DB 39 SerTyrArgThrIleuIleuSerTyrArgTyrArgTyrArgTyrArgTyrArgTyr 58
 QY 550 CAGCAGGCTGAGTGGAGAGACAGGCTCCCGCCCAACCAAGCGGCGAGTGGAGT 609

DB 59 GINGINGLYGINTTTPGLTHRLTYSERPROPHRILYARGINARGSERARGLY 78
 QY 610 CGGCCAGAGTGTGTGCGCAGCGCGGAGAGGGGCCAGCCGACCCAGCAGCAG 669
 DB 79 ARGPROSERGLYGLYALALYARGARGARGARGGLYALAPROALALPROGLINGIN 98
 QY 670 TCAAGCCCGCCAGACTTCTCTGAAAGCAAGTACCTGTGACATCCGCTCCGGGT 729
 DB 99 SERGIUPROALARGPROSERSERGLIUGLYSVALTHRCYASPDILARGLEUARGVAL 118
 QY 730 CGAGAGAGTACTGTGAGCATGGCGCAGCTTGAGCAGGGGGTGGCATCCCGGCGCC 789
 DB 119 ARGALAGLUTRYCYGLIHLISGLYPROALAEUGLINGLYVALALSERARGARGPRO 138
 QY 790 CAGCGCTGCGCGCGCAGCTGAGCAGCTGTTGGCAGGCCACCGAGTCTGCTCAAG 849
 DB 139 GINALAEUVALARGGLNEUASPVALPHEGLYGLINALTHRALVALLEUARGSERARG 158
 QY 850 GACCTGGGCTCTGTGTGTGTGACATCAAGTCTCTGAGCTCTCTATCTGAGCCCTTC 909
 DB 159 ASPLEUGLYSERVALVALCYSPILLEYSPHESERGIULEUSERTYRLEUASPALPHE 178
 QY 910 TGGGGGCACTACTGAGTGGCGCCCTGCTGAGCGGCGGCGGCGTCTCTGACTGAG 969
 DB 179 TRPGLYASPTRYLEUSERTYRLEUVALLEUVALARGLYVALPHELEUTHRGIL 198
 QY 970 GCCCTGCGAGAGCTGTGGCGCGGAGGCTGTTCCCTGCTGCTGCTGAGTGGAGGCT 1029
 DB 199 ALAEUVALAGLVALVALGLYARGGLVALALARGLEUVALSERVALASPGLIUNLA 218
 QY 1030 GACTATGAGGCTGCGCGCGCGCTGCTGCTGATGAGAGAGGAGGCGGCGCGCCG 1089
 DB 219 ASPTRYGLIUNLAGLYARGARGARGLEUVALLEUVALGLIUGLYGLYARGARGPRO 238
 QY 1090 ACAGAGGCTCC 1101
 DB 239 THRGILUALSER 242
 RESULT 14
 AAM41591
 ID AAM41591 standard; protein; 217 AA.
 XX AAM41591;
 AC 22-OCT-2001 (first entry)
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 6522.
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS WO200153312-A1.
 PN 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 PF 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-005598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QJ,
 PI Zhou F, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AA160747.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 2; SEQ ID NO 6522; 10078bp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA158642-AA162213) with neurotrophic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC activation/inhibition activity, chemoclastic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 CC
 XX
 SQ Sequence 217 AA;

Alignment Scores:
 Pred. No.: 5.5e-72 Length: 217
 Score: 1096.50 Matches: 214
 Percent Similarity: 96.86% Conservative: 2
 Best Local Similarity: 95.96% Mismatches: 1
 Query Match: 31.02% Indels: 6
 Gaps: 1

US-10-030-271-3 (1-1883) x AAM41591 (1-217)

QY 37 AACAGTGTGCGCGCGCTGAGCGGAGGATCGAGCGGAGGATCGCGGCTTCTG 96
 DB 1 AAslystrlpneuhargProserProhrgSerHlsrghthrProgluSerGlyhArgValLeu 20
 QY 97 AGCTTGTTCGCGCTCCCTCCCGGGAATGCGGCTATCCGGGTGACCCCGCGCTGC 156
 DB 21 SerleuphargleuProProProgluMetAlaleuSerGlySerThrProAlaProCys 40
 QY 157 TGGGAGGAGATAGTCCCTGACTACTAGCGGATGCTGTCCGCTTACCGGTATGTTGAG 216
 DB 41 TrpGluGluabpGluCysleuAaspTyTrGlyMetleuSerleuHlsAsgMetPheGlu 60
 QY 217 GTGTGGGCGGCGAATCGAGCGAGTGTGAGTCTCGGCGCTTCTGCTGGAGAG 276
 DB 61 ValValGlyGlyGlnleuThrGluCysGlnleuGlnleuLeuAlaPheleuLeuAaspGlu 80
 QY 277 GCTCTGTGCGCGCGCGGAGCTTAGCCCGGAGCCCGAGCGGCTTAGAGCTCTGTGGAG 336
 DB 81 AlaProGlyAlaAlaGlyGlyLeuSerArgAlaArgSerGlyLeuGlySerleuLeuGlu 100
 QY 337 CTGAGGCGCGCGCGGCGAGTGTGCGGAGCAACTGTGGGCTGCTGGGCGCAATCTCTGCGC 336
 DB 101 LeuGluArgArgGlyGlnCysAaspGluSerAaspAaspAaspAaspAaspAaspAasp 120
 QY 397 GTGTGGCGCGCGCGAGCTGTGCGGAGCTGTGCGGAGCGGCGCGGCGAGTGTCT 456
 DB 121 ValLeuAlaArgHlsAaspLeuProHlsleuAlaArgHlsAaspGlyAaspGlyAasp 140
 QY 457 CGAAGCGCTATAGCTATGCGACCTCCAGCTCTTCAAGAAGAGCAGAGGAGTGTGCGCT 516

DB 141 ProGluArgTysSerTyrGlyThrSerSerSerTyrArgThrGluGlySerCysArg 160
 QY 517 CGCGTGTGCGAGCGAGCAAGTCTGCAATTCACAGAGGAGTGTGCGAGCAGGCTCC 576
 DB 161 ArgArgArgGlnSerSerSerSerAlaSerGlnIn-----GlySer 175
 QY 577 CCCCACACAGCGGCGGCGGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 636
 DB 176 ProProThrlyArgGlnArgGlnArgSerArgGlyArgProSerGlyGlyAlaArgArg 195
 QY 637 CGAGAGGCGGCGCGCGCGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 636
 DB 196 ArgArg-GlyProGlnProHlsProSerSerSerGlnSerProProAaspLeuProLeu 215
 QY 697 GGCAGAG 703
 DB 215 sAlaLys 217
 RESULT 15
 AAB94040
 ID AAB94040 standard; protein; 318 AA.
 AC AAB94040;
 AC AAB94040;
 DT 26-JUN-2001 (first entry)
 DT XX
 DE Human protein sequence SEQ ID NO:14195.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS
 PN EP1074617-A2.
 PD 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-ARG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INSTR.
 PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ichii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PS Claim 8; SEQ ID NO 14195; 2537bp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_n2p model

Run on: March 21, 2005, 23:49:20 ; Search time 327.434 Seconds
(without alignments)
5889.702 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 3535
Sequence: 1 aggcgcatacatagagaag.....ctgaaaggcatagctg999 1883

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame-n2p-model -DEV=x1p
-Q=cgnt2_1/USFTO_spool_p/US10030271/runat_21032005.153458.26468/app_query_fasta_1.3150
-DB=uniprot_03 -OPMT=faetan -SUFPIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USERR=US10030271 @CN 1.1 578 @runat 21032005.153458.26468 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: uniprot_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1659	46.9	326	DED2_HUMAN	Q8WTF8 homo sapien
2	1508	42.7	330	DED2_MOUSE	Q8QZV0 mus musculu
3	708	20.0	404	O6DNH2	Q6DNH2 brachydanio
4	700.5	19.8	369	O919M3	Q919M3 brachydanio
5	647.5	18.3	318	DED2_HUMAN	Q75618 homo sapien
6	643.5	18.2	318	DED2_RAT	Q92240 rattus norv
7	641.5	18.1	318	DED2_MOUSE	Q92113 mus musculu
8	530	15.0	168	O8BRM9	Q8BRM9 mus musculu
9	398	11.3	243	O6GNZ8	Q6GNZ8 xenopus lae
10	349.5	9.9	244	O6DHV2	Q6DHV2 brachydanio
11	276.5	8.1	659	O6C708	Q6C708 varrowia li
12	271.5	7.9	585	O41935	O41935 murid herpe
13	269	7.8	660	YHL1_EBV	P03181 Epstein-Bar
14	269	7.8	660	Q777A3	Q777A3 human herpe
15	262.5	7.4	903	O82HR3	Q82HR3 streptomyce
16	259	7.3	660	YHL1_EBV	P03181 Epstein-Bar

17	259	7.3	660	Q777A3	Q777A3 human herpe
18	258.5	7.5	3204	Q6X248	Q6X248 bovine herp
19	253.5	7.4	1163	Q8N6U4	Q8N6U4 homo sapien
20	253.5	7.4	1466	CA13_HUMAN	P02461 homo sapien
21	253	7.4	1461	IE18_PRIVIF	P1675 pseudorabic
22	251	7.3	511	Q95J00	Q95J00 sus scrofa
23	251	7.3	566	Q95J01	Q95J01 sus scrofa
24	251	7.3	816	Q7QWN8	Q7QWN8 giardia lam
25	249.5	7.3	1953	Q9B1T7	Q9B1T7 nephila ina
26	248	7.2	745	Q89X06	Q89X06 brachyribzob
27	248	7.0	1953	Q9B1T7	Q9B1T7 nephila ina
28	247.5	7.2	676	Q95J09	Q95J09 sus scrofa
29	247.5	7.2	1447	Q95I91	Q95I91 xenopus lae
30	247	7.2	890	Q770B7	Q770B7 alvinella p
31	245.5	7.2	1449	Q802B5	Q802B5 xenopus lae
32	245	7.1	1446	IE18_PRIVKA	P33479 pseudorabic
33	243.5	7.1	1083	IE18_HUMAN	Q00268 homo sapien
34	243	7.1	496	Q708L8	Q708L8 synechococc
35	240.5	6.8	1070	Q960G5	Q96J95 homo sapien
36	238	6.9	608	Q9Q5K9	Q9Q5K9 herpesvirus
37	238	6.7	1453	CA11_CHICK	P02457 gallus galli
38	237	6.9	2157	Q9Z1R1	Q9Z1R1 mus musculus
39	237	6.9	2158	Q7TSC1	Q7TSC1 mus musculus
40	236.5	6.9	1454	Q8ULT63	Q8J163 suid herpes
41	236	6.9	1447	Q26U13	Q26U13 brachydanio
42	235	6.8	1414	Q26634	Q26634 strongyloce
43	235	6.8	1464	Q8BLW4	Q8BLW4 mus musculus
44	235	6.8	2157	Q95875	Q95875 homo sapien
45	234.5	6.8	1449	Q6PE19	Q6PE19 brachydanio

ALIGNMENTS

RESULT 1
DED2_HUMAN STANDARD; PRT; 326 AA.
ID DED2_HUMAN
AC Q8WTF8; Q8NBR2; Q8NBS1; Q8TAAB; Q96D35;
DT 05-JUN-2004 (Rel. 44, Created)
DT 05-JUN-2004 (Rel. 44, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE DED2_HUMAN death effector domain-containing protein 2 (DED-containing protein FLAME-3).
DE protein FLAME-3).
GN Name=DED2; Synonyms=FLAME3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND INTERACTION WITH CASP8.
RP MEDLINE=21850646; PubMed=11741985; DOI=10.1074/jbc.M10749700;
RA Roth W., Steiner-Liewen F., Pawlowski K., Godzik A., Reed J.C.;
RT "Identification and characterization of DED2, a death effector domain-containing protein.";
RT J. Biol. Chem. 277:7501-7508(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, SUBCELLULAR LOCATION, AND INTERACTIONS WITH CASP8 AND GTRF33.
RP MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T., Alnemri E.S.;
RT "Death effector domain-containing proteins DED2 and FLAME-3 form nuclear complexes with the TRIFC102 subunit of human transcription factor TRIF.";
RT Cell Death Differ. 9:439-447(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RP MEDLINE=22202288; PubMed=12235123; DOI=10.1093/jcb.200112124;
RA Lee J.C., Schickling O., Stegh A.H., Oshima R.G., Dinsdale D., Cohen G.M., Peter M.E.;
RT "DED2 regulates degradation of intermediate filaments during apoptosis.";
RT J. Cell Biol. 158:1051-1066(2002).

[4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Tetacarcinoma;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Ito K., Kawai T., Isono Y., Nakamura Y.,
 RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi K., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ihibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hottu T.,
 RA Kusano Y., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togaiya S., Komai F., Hara R., Takuchi K., Arita M.,
 RA Imose N., Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsumura H., Ichihara T., Shihara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamaguchi M., Komiyama K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hitao M., Omori Y.,
 RA Kawabata S., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi I., Maeno Y., Yamagawa R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs";
 RN Nat. Genet. 36:40-45(2004).
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Brain; and Lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallory S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maria W.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP INTERACTIONS WITH CASP8 AND CASP10.
 RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
 RA Alcivar A., Hu S., Tang J., Yang X.;
 RT "DEDD and DEDD2 associate with caspase-8/10 and signal cell death.";
 RN Oncogene 22:291-297(2003).
 RP - FUNCTION: May play a critical role in death receptor-induced
 RP apoptosis and may target CASP8 and CASP10 to the nucleus. May
 RP regulate degradation of intermediate filaments during apoptosis.
 RP May play a role in the general transcription machinery in the
 RP nucleus and might be an important regulator of the activity of
 RP GTF3C3.
 RP - SUBUNIT: Interacts with CASP8, CASP10 and GTF3C3. Homodimerizes
 RP and heterodimerizes with DEDD.

CC -1- SUBCELLULAR LOCATION: Nuclear; accumulated in subnuclear
 CC structures resembling nucleoli.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8WXR8-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8WXR8-2; Sequence=VSP_010312;
 CC -1- TISSUE SPECIFICITY: Expressed in most tissues. High levels were
 CC found in liver, kidney, heart, ovary, spleen, testes, skeletal
 CC muscle and peripheral blood leukocytes. Expression was absent or
 CC low in colon and small intestine. Expression is relatively high in
 CC the tumor cell lines chronic myelogenous leukemia K-562 and the
 CC colorectal adenocarcinoma SW480. Expression is moderate in the
 CC cervical carcinoma HeLa, the Burkitt's lymphoma Raji, the lung
 CC carcinoma A549, and the melanoma G361. In contrast, two leukemia
 CC cell lines, HL-60 (promyelocytic leukemia) and MOLT-4
 CC (lymphoblastic leukemia), show relatively low levels.
 CC -1- DOMAIN: Interacts with CASP8 and CASP10 are mediated by the DED
 CC domain.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -1- CAUTION: Ref.5 (AAH3372) sequence differs from that shown due to
 CC a frameshift in position 186.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL; AF43591; AAL4820.1; -;
 CC DR EMBL; AF457575; AAM10835.1; -;
 CC DR EMBL; AY125488; AAM95240.1; -;
 CC DR EMBL; AK075328; BAC11551.1; -;
 CC DR EMBL; BC013372; AAH3372.2; ALT_FRAME.
 CC DR EMBL; BC027930; AAH27930.1; -;
 CC DR Genew; HGNC:24450; DEDD2.
 CC DR H-InvDB; HIX0015171; -;
 CC DR InterPro; IPR011029; DEATH_like.
 CC DR InterPro; IPR001875; DED.
 CC DR Pfam; PF01335; DED_1.
 CC DR SMART; SM00031; DED; 1.
 CC DR PROSITE; PS0168; DED; 1.
 CC KW Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
 CC Transcription regulation.
 CC FT DOMAIN 25 104 DED.
 CC FT DOMAIN 104 109 Nuclear localization signal (Potential).
 CC FT DOMAIN 155 173 Bipartite nuclear localization signal
 CC (Potential).
 CC FT VARSPPLIC 145 149 Missing (in isoform 2).
 CC FT CONFLICT 27 27 /FTID=VSP_010312.
 CC FT CONFLICT 56 56 A -> N (in Ref. 2).
 CC FT CONFLICT 79 79 D -> G (in Ref. 4).
 CC FT CONFLICT 207 207 C -> R (in Ref. 5).
 CC FT CONFLICT 230 230 Missing (in Ref. 5; AAH3372).
 CC FT SEQUENCE 326 AA; 36178 MW; 3F7B0B307CC870CD CRC64;
 CC SQ
 CC Alignment Scores:
 CC Pred. No.: 5,49e-79 Length: 326
 CC Score: 1659.00 Matches: 325
 CC Percent Similarity: 99.69% Conservative: 0
 CC Best Local Similarity: 99.69% Mismatches: 1
 CC Query Match: 46.93% Indels: 0
 CC DB: 1 Gaps: 0
 CC US-10-030-271-3 (1-1883) x DEDD2_HUMAN (1-326)
 CC QY 124 ATGGCGCTATCCGGGCGACCCCGGCGCTGCGAGAGATGAGCTGACACTAC 183
 CC DB 1 MetAlaLeuSerGlySerThrProAlaProCySTPGLUGLInuSpGLuCyLeuAspTyr 20

QY 184 TACGGATGCTGTCGCTTCAACCGTATGTTGAGGTGTGGCGGCAACGAGTGC 243
 DB 21 TycilymetleusSerleuH1sArgMetPheGluValaGlyGlyInleuthrGlucys 40
 QY 244 GAGCTGAGAGCTCCCTGCGCTTCTGCTGAGAGAGCTCTCGCGCGCGGAGGTTAGCC 303
 DB 41 GluleuGluInleuEuAlaPheLeuLeuAspGluLaProGlyAlaAlaGlyGlyLeuAla 60
 QY 304 CGGCGCCGAGCGGCTAGAGAGCTCTGCTGAGAGCTGAGAGCGCGCGGAGTGGCGAG 363
 DB 61 ArgAlaArgSerGlyLeuGluInleuLeuGluInleuGluInleuGluInleuGluInleu 80
 QY 364 AGCAACCTGGCGCTGCTGGCGCAACTCTGGCGGTGTGGCGCGCAAGACTGCTCCG 423
 DB 81 SerAsnleuArgleuLeuGlyGlyInleuLeuArgValleuAlaArgHisAspLeuPro 100
 QY 424 CACTGGCGCGGAGCGGCGCGCGCGCTGCTCTCAAGAGCTTACTATGAGACTCC 483
 DB 101 HisleuAlaArglySerArgArgProValSerProGluArglySerTyrglyThrSer 120
 QY 484 AGCTCTCAAGAGAGAGAGAGAGAGTGCCTGCGCGCTGCGCGAGTCAAGAGTTCGCA 543
 DB 121 SerSerSerlySerArgThrGlyGlySerCyArgArgArgArgGlnSerSerSerAla 140
 QY 544 AATTCTCAGAGAGGCTCAGTGGAGAGAGAGCTCCCGCCCAACCAAGCGAGCGGAGT 603
 DB 141 AsnSerGlnGlnGlyInleuGlyInleuGlySerProThrlySerArgAlaArgSer 160
 QY 604 CGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
 DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgArgArgArgArgArgArg 180
 QY 664 CAGAGAGAGAGCGCGCGAGAGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723
 DB 181 GlnGlnSerGluProAlaArgProSerSerGlyGlyValThrCyAspIleArgLeu 200
 QY 724 CGGAGTTCAGAGAGAGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
 DB 201 ArgValaArgAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 220
 QY 784 CGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
 DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 844 TCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
 DB 241 SerArgAspLeuGlySerValValCyAspIleLeuSerGlyLeuSerTyrglyLeuAsp 260
 QY 904 GCTTCTGAGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
 DB 261 AlaPheArgGlyAspTyrglyLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 964 ACTGAG 1023
 DB 281 ThrGluAlaLeuArgGlyAlaValaGlyArgGlyAlaValaArgLeuLeuValaSerValaAsp 300
 QY 1024 GAGGCTACTATGAG 1083
 DB 301 GluAlaAspTyrglyAlaGlyArgArgArgArgLeuLeuLeuGluGluGlyGlyArg 320
 QY 1084 CGGCGCGAGAGAGAGCTCC 1101
 DB 321 ArgProThrArgLeuAlaSer 326
 RESULT 2
 DED2_MOUSE STANDARD; PRF; 330 AA.
 ID DED2_MOUSE
 AC 08QZV0; 08JZV1;
 DT 05-JUN-2004 (Rel. 44, Created)
 DT 05-JUN-2004 (Rel. 44, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE DNA-binding death effector domain-containing protein 2 (DED-containing

DE protein FLAME-3).
 GN Name=Ded2; Synonyms=Flame3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6J;
 RX MEDLINE=2388257; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
 RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T.,
 RA Alnemri E.S.;
 RT "Death effector domain-containing proteins DEDD and FLAME-3 form
 RT nuclear complexes with the p110C12 subunit of human transcription
 RT factor IIIC."
 RL Cell Death Differ. 9:439-447(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
 RA Alcivar A., Hu S., Tang J., Yang X.;
 RT "DEDD and DEDD2 associate with caspase-8/10 and signal cell death."
 RL Oncogene 22:291-297(2003).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Retina;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Haile F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Lomelianno N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: May play a critical role in death receptor-induced
 CC apoptosis and may target Casp8 and Casp10 to the nucleus. May
 CC regulate degradation of intermediate filaments during apoptosis.
 CC May play a role in the general transcription machinery in the
 CC nucleus and might be an important regulator of the activity of
 CC GPR33.
 CC -1- SUBUNIT: Interacts with CASP8, CASP10 and GPR33. Homodimerizes
 CC and heterodimerizes with DEDD (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear, accumulated in subnuclear
 CC structures resembling nucleoli (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=08QZV0-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=08QZV0-2; Sequence=VSP 010313, VSP 010314;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Expression is high in liver, heart, kidney,
 CC and testis but low in brain, spleen, lung, and skeleton muscle.
 CC -1- DOMAIN: Interactions with Casp8 and Casp10 are mediated by the DED
 CC domain (By similarity).
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL AF457576; AAM10836.1; -
 DR EMBL AF543541; AAN33179.1; -
 DR EMBL BC037043; AAN37043.1; -
 DR MGD; MGI:1914629; 2410050E1R1K.
 DR InterPro; IPRO11029; DEATH_like.
 DR InterPro; IPRO1875; DED.
 DR Pfam; PF01335; DED; 1.
 DR SMART; SM00031; DED; 1.
 DR PROSITE; PSS0168; DED; 1.
 DR Alternative splicing; Apoptosis; Nuclear protein;
 KW Transcription regulation.
 FT DOMAIN 25 104 DED.
 FT DOMAIN 104 109 Nuclear localization signal (Potential).
 FT DOMAIN 156 174 Bipartite nuclear localization signal (Potential).
 FT VARSPLIC 1 171 Missing (in isoform 2).
 FT VARSPLIC 172 200 /FTID=VSP_010313.
 FT /RRAGLAASQOHQOHELGRPSSEKRVTC -> MGKQAGRG
 FT EVTCPEPTRAMOSMDSLS (in isoform 2).
 FT /FTID=VSP_010314.
 SQ SEQUENCE 330 AA; 36786 MW; 889BC4F9E01304B0 CRC64;

Alignment Scores:

Pred. No.: 4.38e-71 Length: 330
 Score: 1508.00 Matches: 303
 Percent Similarity: 93.33% Conservative: 5
 Best Local Similarity: 91.82% Mismatches: 18
 Query Match: 42.66% Indels: 4
 DB: 1 Gaps: 2

US-10-030-271-3 (1-1883) x DED_MOUSE (1-330)

QY 124 ATGGCGCTTTCGGGCTGACCCCGCGCTGTGGAGAGAGATGACTGCTGACTAC 183
 DB 1 MetAlaLeuSerGlySerThrProAlaProSerTrpGluGluAspIuGluLeuMetGlyTyr 20
 QY 184 TTCGGGATCTGTCGCTTTCGCTGATGAGGCTCTGCGCGCGCGGACCTGAGCTGC 243
 DB 21 TTTGlyMetLeuSerLeuHisArgMetPheGluValaGlyGlyIleuMetGluGlyCys 40
 QY 244 GAGCTGAGCTCTGCGCTTTCGCTGATGAGGCTCTGCGCGCGCGGACCTGAGCTGC 303
 DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAlaPheGluValaProGlyAlaProGlyIleuAla 60
 QY 304 CGGCGCGCGAGCGGCTTTCGCTGAGCTGAGGCTGAGGCGCGCGGACCTGAGCTGC 363
 DB 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluValaArgGlyIleuGlyCysArgGlu 80
 QY 364 AGGAACCTCGGCTGTGGGGGCACTCTGCGGCTGCGCGCGCGGACCTGAGCTGC 423
 DB 81 SerAsnLeuArgLeuLeuSerGlnLeuLeuArgValaLeuAlaArgHisAspLeuLeuPro 100
 QY 424 CACCTGCGCGAGCGGCGCGCGGCTTTCGCTGAGGCTGAGGCTGAGGCGCGGACCTGC 480
 DB 101 HisLeuAlaArgIleArgIleArgIleProValaSerProGluArgIleArgIleArgIlePro 120
 QY 481 TTCAGCTCTTCAAG 540
 DB 121 SerSerSerSerIleArgIleArgIleArgIleArgIleArgIleArgIleArgIleArgIle 140
 QY 541 GCAAAATTCAG 600
 DB 141 SerAspSerProGlnArgIleArgIleArgIleArgIleArgIleArgIleArgIleArgIle 160
 QY 601 AGTCGCGCGCGCGAGTGGTGGTGCAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 161 SerArgGlyArgProSerSerGlyAlaArgGlnArgIleArgIleArgIleAlaAlaSer 180
 QY 661 -----CAGCAGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 711

DB 181 GlnGlnHisGlnGlnHisGlnGlnIleuGluValaArgProSerSerGlnGlyValaThrCys 200
 QY 712 GACATCCGCGCTCCGGGCTTCGAGCAGAGTACTGCGAGCATGGGCGACCTTGGAGAGAGGC 771
 DB 201 AspIleArgLeuArgValaArgAlaGluTyrCysGlnHisGlyProAlaLeuGluGlnGly 220
 QY 772 GTGGCATCCCGGCGCGCGCGCGGCTGCGCGCGCGGACCTGAGCTGTTGGGCGAGCCACC 831
 DB 221 ValAlaSerArgArgProGlnAlaLeuAlaArgGlnLeuAspValaPheGlyGlnAlaThr 240
 QY 832 GCAGTGTGCGGCTCAGAGGACCTGCGGCTGCTGTTGGATCATGACTTCAGAGCTC 891
 DB 241 AlaValLeuArgSerArgAspLeuIleValaValCysAspIleuArgPheSerGluLeu 260
 QY 892 TCCTATCTGAGAGCCCTTTCGCGGCGACTTACCTGAGTGGCGCCCTGCGAGGCGCTGCGG 951
 DB 261 SerTyrLeuAspAlaPheThrGluAlaLeuArgGluAlaValaGlyArgGluAlaValaArgLeu 280
 QY 952 GCGGTGTTCTCAGCTGAGGCGCTGCGAGAGCTGTGGGCGCGGAGAGCTGTGCTGCTG 1011
 DB 281 GlyValaPheLeuThrGluAlaLeuArgGluAlaValaGlyArgGluAlaValaArgLeu 300
 QY 1012 GTCAGTGTGATGAGGCTGACTATGAGGCTGCGCGCGCGCTGCTGCTGATGAGAGAG 1071
 DB 301 ValSerValaAspGluAlaAspIleArgIleArgIleArgIleArgIleArgIleArgIle 320
 QY 1072 GAAGGGGGGCGCGCGCGCGCGAGAGGCTTCC 1101
 DB 321 GluGlyGlyArgArgGlyThrGluAlaSer 330
 RESULT 3
 Q6DHN2 PRELIMINARY; PRT; 404 AA.
 AC Q6DHN2;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DE 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 Zgc:92202.
 GN Name=zgc:92202;
 OS Brachydanio rerio (zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 ON NCBI_TaxID=7955;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner I., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Strausberg R.,

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC075935; AAH75935.1; -

DR GO; GO:0005515; F:protein binding; IEA.

DR GO; GO:0042981; P:regulation of apoptosis; IEA.

DR InterPro; IPR011029; DEATH_Like.

DR InterPro; IPR001875; DED.

DR Pfam; PF01335; DED; 1.

DR SMART; SM00031; DED; 1.

DR PROSITE; PS50168; DED; 1.

SO SEQUENCE 404 AA; 4488 MW; 3C949DB3B07B81A CRC64;

Alignment Scores:

Pred. No.:	3,19e-29	Length:	404
Score:	708.00	Matches:	163
Percent Similarity:	60.32%	Conservative:	65
Best Local Similarity:	43.12%	Mismatches:	106
Query Match:	20.03%	Indels:	44
DB:	2	Gaps:	10

US-10-030-271-3 (1-1883) x Q6DHN2 (1-404)

```

QY 10 AATGAGAGAGTGCAGAAAGATCCAAAGATGGCTGGCGCGCCGAGATCAT 69
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 37 SerArgSerGlyLeuSerSerProArgGlyTyrValAlaIaIaSer---SerSerGly 55
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 CGAGCGCCAGAAATGCGCGCGGTTCTGAGCTTGTTCGCGCTCCCTCCCGCGGAGATGCG 129
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 56 ThrThrHisGlyAanSerLeuAlaLeuSerLeuSerLeuAlaProAlaSerSerSer 75
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 CTATCCGGGTGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 171
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 AsnSerSerSerThrArgArgProAlaSerGlyArgValGluProGluGluGlu 95
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 172 TGCGTGAATCTCTGCGGAGTGTGCTTCACCGTATGTTGAGGTGTGGCGGCGCA 231
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 96 ValAspAspAlaTyrGlyLeuTyrSerLeuHisArgPheSerPheAspIleValGlyAlaGln 115
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 232 CTGACCGAGTGCAGCTGCAGACTCTGCGCTCTCTCTG-----CTGAGTGAAGCT 279
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 116 LeuThrHisArgAspValArgValLeuSerPheLeuPheValAlaIaIaSerGluTyr 135
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 280 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 339
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 -----GluArgGlyGlyIle-----ArgSerGlyArgAspPheLeuLeuAlaLeu 150
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 340 GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 399
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 GluArgGlnGlyArgCyAspGluThrAsnPheArgHisValLeuGlnLeuLeuArgIle 170
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 CTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 459
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 IleThrArgHisAspLeuLeuProTyrValThrLeuArgValArgGlnThrValCyAspPro 190
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 GAACGCTATAGCTATGGCACTCCAGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 Asp-----ProValAspTyrTyrLeuGluGluThrSerValArg 203
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 CGTGGGAGTGAACAGATTCTGCAAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 204 TyrValSerProArgGlyThrAlaAspIleGlnGlnGlnThrProHisArgArgThrGly 223
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 574 -----TCCCGCCCAACCAAGCGCGAG-----CGCGGAGTGTGCG 606
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 224 ProGlnProLeuIleCyAspProProSerGlyProGlnValCyAspProProArgAlaIys 243
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 GCGCGCGCGCGAGTGTGTGCGAGACGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 244 ProAlaProProProProSerArgIleArgIleArgIleArgIleArgIleArgIleArgIle 263
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 667 CAGTCAGAGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 264 Glu-----LysGlnThrCyAspIleArgLeuArg 273

```

```

QY 727 GTTGACCAAGATCTGCAGCATGGCGACGCTTGAGGAGAGGCGCTGSCATCCCGCGG 786
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 274 ValArgIleGluTyrCyAspGlnHisGluSerAlaLeuGlnGlnValPheSerAlaIys 293
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 787 CCCGAGCGCGTGGCGCGCGAGCTGAGAGTGTGGGAGCGCCACCGAGTGTGGCTCA 846
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 294 GlnGluAlaLeuGluArgGlnPheGluArgPheAsnGlnAlaThrIleLeuLysSer 313
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 847 AGGAGCTGGAGCTGTGTGTGTCATCAAGTTCAGAGCTCTCTTATGAGAGCC 906
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 314 ArgAspLeuGlySerIleIleCyAspIleArgPheSerGluLeuThrTyrLeuAspAla 333
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 907 TTCTGGGCGCATCTAGCTGAGCGCGCTGCTGAGCGCGCGCGCGCTGTCTGACT 966
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 334 PheThrArgAspTyrIleAlaGlnGlySerLeuLeuGlnAlaLeuGlyValPheIleThr 353
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 967 GAGCGCTGGAGAGAGCTGTGTGTGGCGCGCGAGAGCTTGTGCTGTGTGAGTGAAGAG 1026
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 354 AspSerLeuGlyGlnAlaValGlyHisGlyAlaIleLysLeuLeuValAlaSerValAspGlu 373
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1027 GCTGACTATGAGCTGGCGCGCGCGCGCGCTGTGCTGATGAGAGAGAGAGAGAGAGAG 1080
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 374 GluAspTyrGlnAlaGlyArgArgLysLeuLeuArgAsnLeuValAlaGlyGly 391
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
ID Q919M3 PRELIMINARY; PRT; 369 AA.
AC Q919M3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Deduc1.
GN Name=Deduc1;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish."
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF232226; AAR6963.1; -.
DR ZFIN; ZDB-GENE-000616-2; deduc1.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro; IPR011029; DEATH_Like.
DR InterPro; IPR001875; DED.
DR Pfam; PF01335; DED; 1.
DR PROSITE; PS50168; DED; 1.
SO SEQUENCE 369 AA; 42244 MW; 73B09E9E17ECC247 CRC64;

Alignment Scores:
Pred. No.: 7.87e-29 Length: 369
Score: 700.50 Matches: 161
Percent Similarity: 56.70% Conservative: 38
Best Local Similarity: 45.87% Mismatches: 93
Query Match: 19.82% Indels: 59
DB: 2 Gaps: 7

US-10-030-271-3 (1-1883) x Q919M3 (1-369)
QY 157 TGGAGAGAGATGAGTCTGAGCTATGAGGAGATGCTGTGCTTACCGGTATGTTGAG 216
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 13 TrpGluGluThrGlnGlyLeuSerTyrTyrGluThrLeuSerLeuHisGluIlePheGlu 32
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 GTGTGGCGCGGCACTGAGACCGAG---TGGAGACTGAGACTCTGTGCGCTTGTGCTGAT 273
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 33 IleValGlySerGlnLeuThrGlnThrCyAspGlyGlyGlu---ValAlaPheLeuLeuAsp 51
DB   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 274 GAGGCT-----CCT 282
 DB 52 GluThrTyrProGlyLysHisProLeuAspProGluGlyTyrThrGluAspLeuProPro 71
 QY 283 GGGCGCGCGGGA----- 294
 DB 72 GlyProAspGlySerProGlnAlaSerThrProCysProArgLeuLeuLysSerTyrGln 91
 QY 295 -----GGCTTAGCCCGGCGCGCAGCGCG 318
 DB 92 ArgMetGlnProGlnLysGluGlyCysSerIleAlaSerArgHisArgProLysSerGly 111
 QY 319 CTAGAGCTCTCTGTGAGCTGAGCGCCCGCGGCGAGCTGCGCGAGCAACTCGCGCTG 378
 DB 112 ValGluLeuLeuLeuLysGluLeuGluArgGlyTyrLeuSerAspValAlaSerLeuAspPro 131
 QY 379 CTGGGGGCACTCTGCGCGCGCGCGCGCGAGCACTGCTGCGCGAGCGCGCGCGAG 438
 DB 132 LeuLeuGlnLeuLeuArgIleLeuThrArgHisAspValLeuProPheValSerGlnLys 151
 QY 439 CGGCGCGCGCGAGTGTCTCCAGAACGCTATAGCTATAGCACTTCCTTCAAGAGAG 498
 DB 152 LysArgArgThrValSerProGluArgGlnLysIleAspTyrProGluValAspPheArg 171
 QY 499 ACAGAGGAGTACCTGCGCGCGCGCGCGAGCACTTCGCAAAATTCACAGCGGT 558
 DB 172 GlnAspArgGluValGlySerAsnThrAsnIleProSerPheGluAsnThrGlnAspHis 191
 QY 559 CAGTGGGAGACAGGCTCC-----CCCCAACCAAGCGGAGCGCGCGAGTCCGG 606
 DB 192 HisThrArgAlaGlySerGlySerMetThrSerAlaSerSerAsnArgGlyAspArg 211
 QY 607 GGGCGCGCGCGAGTGTGTGCGAGACCGCGCGAGAGGAGGCGCGCGCGCGCGAG 666
 DB 212 GlyArgGlyHisHisThrSerArgLysSerArg-----GlyProProGluIleGlnPro 229
 QY 667 CAGTCAGAGCGCGCGAGCTTCTTGAAGGCAAAAGTGCATCTGACATCCGGCTCCGG 726
 DB 230 GlnSerThrPro-----AsnIleValThrCysAspIleLeuArgLeuArg 243
 QY 727 GTTTCGAGAGAGTACTCGAGCATGGGCGCAGCTTGAGAGAGGCGCGTGGCATCCGCGCG 786
 DB 244 ValArgAlaGluTyrSerGlnHisGluSerAlaLeuArgGlyGlyPheProSerAspLys 263
 QY 787 CCCCAGCGCTGCGCGCGAGCTGAGCATGTTTGGCGAGCGCGCAGCGAGTGTCCGCTCA 846
 DB 264 ProGlnProLeuGlnArgGlnPheGlnLeuPheSerArgAlaSerLeuLeuLeuArgThr 283
 QY 847 AGGAGACTGGGCTGTGTGTGTGATCATCAATTCACAGAGCTTCTTCTGATCGAGCGCC 906
 DB 284 ArgAspLeuGlySerIleValCysAspIleLysPheSerLysLeuThrAsnLeuAspThr 303
 QY 907 TTCTGCGGGGAGTACTGAGTGGCGCGCTGCTGAGAGCGCTGCGGGCGGTTCCTGAGCT 966
 DB 304 PheThrProAspLysThrSerGlyAlaLeuLeuLysAlaLeuLysGlyValPheIleThr 323
 QY 967 GAGGCCCTCGAGAGGCTGTGTGCGCGCGAGGCTGTTTGCCTGCTGATCGATGAG 1026
 DB 324 AspSerLeuIleArgAlaIleArgGlnLysGlyValArgLeuLeuValSerValAspGln 343
 QY 1027 GCTGACTATGAGGCTGCGCGCGCGCGCTGTTG 1059
 DB 344 TyrAspTyrGluGluGlyArgLysLeuLeuIle 354
 RESULT 5
 DEDD_HUMAN STANDARD; PRT; 318 AA.
 ID 075618; 060737;
 AC 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Death effector domain-containing protein (Death effector domain-containing testicular molecule) (DEDProl) (FLNED-1) (KE05).

GN Name=DEDD; Synonyms=DEDProl, DEFT;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=98447599; PubMed=9774341; DOI=10.1093/emboj/17.20.5974;
 RA Stegh A.H., Schickling O., Ehret A., Scalfidi C., Peterhansel C.,
 RA Hofmann T.G., Grunmt I., Kramer P.H., Peter M.E.;
 RT "DEDD, a novel death effector domain-containing protein, targeted to
 RT the nucleolus.";
 RL EMBO J. 17:5974-5986(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
 RC TISSUE=Testis;
 RX MEDLINE=99049260; PubMed=9832420; DOI=10.1210/en.139.12.4839;
 RA Leo C.P., Hsu S.Y., McGee E.A., Salanova M., Haneh A.J.W.;
 RT "DEFT, a novel death effector domain-containing molecule predominantly
 RT expressed in testicular germ cells.";
 RL Endocrinology 139:4839-4848(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Pan G.;
 RT "FLNED-1, a novel molecule with a DED-like domain.";
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Thome M., Tschopp J.;
 RT "DEDProl, a novel DED-containing protein.";
 RL Submitted (Oct-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Dendritic cell;
 RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;
 RT "A novel gene from human dendritic cell.";
 RL Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Bone marrow, and Placenta;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grove L.H., Derge J.G.,
 RA Klausberg R.D., Feltgen D., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.U., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Myers R.M.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzyzinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP FUNCTION, AND INTERACTIONS WITH KRT8, KRT18 AND CASP3.
 RX MEDLINE=22220288; PubMed=12235123; DOI=10.1083/jcb.200111214;
 RA Lee J.C., Schickling O., Stegh A.H., Oshima R.G., Dindale D.,
 RA Cohen G.M., Peter M.E.;
 RT "DEDD regulates degradation of intermediate filaments during
 RT apoptosis.";
 RL J. Cell Biol. 158:1051-1066(2002).
 RN [8]
 RP INTERACTION WITH GTF3C3.
 RX MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
 RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T.,

RA Alnemrt E.S.;
 RT "Death effector domain-containing proteins DEDD and FLAME-3 form
 RT nuclear complexes with the p71IIC102 subunit of human transcription
 RT factor IIC.";
 RL Cell Death Differ. 9:439-447(2002).
 RN [9]
 RP INTERACTIONS WITH CASP8 AND CASP10.
 RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
 RA Alciivar A., Hu S., Tang J., Yang X.;
 RT "DEDD and DEDD2 associate with caspase-8/10 and signal cell death.";
 RL Oncogene 22:291-297(2003).
 CC -1- FUNCTION: A scaffold protein that directs CASP3 to certain
 CC substrates and facilitates their ordered degradation during
 CC apoptosis. May also play a role in mediating CASP3 cleavage of
 CC KR18. Regulates degradation of intermediate filaments during
 CC apoptosis. May play a role in the general transcription machinery
 CC in the nucleus and might be an important regulator of the activity
 CC of GTF3C3. Inhibits DNA transcription in vitro (by similarity).
 CC -1- SUBUNIT: Interacts with CASP8, CASP10, KR18, CASP3 and
 CC FADD. Homodimerizes and heterodimerizes with DEDD2.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
 CC during CD95-mediated apoptosis where it is localized in the
 CC nucleolus (by similarity). Following apoptosis induction, the mono
 CC and/or dimerization form increases and forms filamentous
 CC structures that colocalize with KR8 and KR18 intermediate
 CC filament network in simple epithelial cells.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=2;
 CC Name=1;
 CC IsoId=O75618-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O75618-2; Sequence=VSP_003846;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
 CC testis.
 CC -1- PTM: Exists predominantly in a mono- or dimerized form.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF083236; AAC33105.1; -
 CC EMBL: AF100341; AAD16414.1; -
 CC EMBL: AF043733; AAC80280.1; -
 CC EMBL: AJ010973; CA09445.1; -
 CC EMBL: AF064605; AAC17110.3; -
 CC EMBL: BC016724; AAH16724.1; -
 CC EMBL: BC013910; AAH13910.1; -
 CC GeneW: HGNC:2755; DEDD.
 CC H-InvDB: HIK0001231; -
 CC MIM: 606841; -
 CC DR GO: GO:0005737; C:cytoplasm; ISS.
 CC DR GO: GO:0005730; C:nucleolus; ISS.
 CC DR GO: GO:0003677; F:DNA binding; ISS.
 CC DR GO: GO:0006912; P:induction of apoptosis; ISS.
 CC DR GO: GO:0008625; P:induction of apoptosis via death domain rec. .; TAS.
 CC DR GO: GO:0016481; P:negative regulation of transcription; ISS.
 CC DR GO: GO:0007283; P:spermatogenesis; TAS.
 CC DR InterPro: IPR011029; DEATH_Like.
 CC DR InterPro: IPR001875; DED.
 CC DR Pfam: PF01335; DED; 1.
 CC DR SMART: SM00031; DED; 1.
 CC DR PROSITE: PS00168; DED; 1.
 CC KM Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
 CC Represor; Transcription regulation.
 CC FT DOMAIN 25 103 DED.
 CC FT VARSPLIC 194 194 D -> GBEIOGFORMSHLEGEYKELLGHMAVVAIQY (1n
 CC isoform 2).

FT FT CONFLICT 13 13 /FTId=VSP_003846.
 SQ SEQUENCE 318 AA; 36794 MW; FF9D5FF9B61F6BB6 CRC64;
 P -> L (in Ref. 5).
 Alignment Scores:
 Pred. No.: 4.68e-26 Length: 318
 Score: 647.50 Matches: 141
 Percent Similarity: 63.55% Conservative: 56
 Best Local Similarity: 45.48% Mismatches: 92
 Query Match: 18.32% Indels: 21
 DB: 1 Gaps: 6
 US-10-030-271-3 (1-1883) x DEDD_HUMAN (1-318)
 QY 157 TGGAGAGAGATGATGCTGAGTCACTACGAGGATGCTGCTTCAACCTGATGTCAG 216
 ||| |||:
 Db 12 TTPProGluGluInuIsGluGluGluInuIsGluLeuTySerLeuInuIsArgMetPheAsp 31
 QY 217 GTGGTGGCGGGGCACTGACCGAGTGCAGTGCAGTGCCTGACCTTCTG--CTGAT 273
 ||| |||:
 Db 32 IlleValGlyThrHisLeuThrHisArgAspValArgValLeuSerPheLeuPheValAsp 51
 QY 274 GAGGCTCTGGCCGCCCGGAGGCTTACGCCGCGGAGCGGCGCTTACAGCTCTGCTG 333
 ||| |||:
 Db 52 ValIleAspAspHisGluArgGlyLeu----IleArgAsnGlyArgAspPheLeuLeu 69
 QY 334 GAGCTGAGCGCGCGGCGGAGTGCAGGAGCACTGCGGCTGCTGCGGCAACTCTCTG 393
 ||| |||:
 Db 70 AlaleuGluArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 89
 QY 394 CGCGTGTGCGCCCGGAGCACTGCTGCGCGGAGTGCAGGCGGCGGCGGCGGCGGAGT 453
 ||| |||:
 Db 90 ArgIleIleThrArgHisAspLeuLeuProTyrValThrLeuLysArgArgAlaVal 109
 QY 454 TCTCCA-----GAGCCTTACGCTATGCACTCC-----AGCTTTCAAAG 455
 ||| |||:
 Db 110 CysProAspLeuValAspLeuTySerLeuGluGluThrSerIleArgTyrValThrProArg 129
 QY 496 AGGACAGAGGATGATGCTGCGTGCCTGCGGCGGAGTCAAGTCTGCAATTCTCAGCAG 555
 ||| |||:
 Db 130 AlaleuSerAspProGluProArgProArgProArgProArgProArgProArgProArg 149
 QY 556 GGTTCAG-----TGGAGNAGAGGCTCCCGGCAACGAGCGGAGCGGCGGAGTCCGAGC 609
 ||| |||:
 Db 150 ProValValLysCysProThrSerGlyProGluMetCysSerLeuLysArgProAlaArgGly 169
 QY 610 CGGCCAGATGCTGCTGTCAGACGCGCGGAGGAGGAGGCGGAGCGGAGCGGAGCGGAGC 669
 ||| |||:
 Db 170 ArgAlaThrLeuGluGluArgGluArgGluArgGluArgGluArgGluArgGluArg 180
 QY 670 TCAGAGCGCGCGGAGCACTTCTCTGAAAGGCAAGTCACTGACATCCGGCTCCGGGTT 729
 ||| |||:
 Db 181 --LysSerValThrProAspProLysGluLysGluThrCysAspIleAspLeuArgVal 199
 QY 730 CGAGCAGAGTACTGTCAGACGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 789
 ||| |||:
 Db 200 ArgAlaGluTyrCysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 219
 QY 790 CAGGCGTGGCGGCGGAGTGCAGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 849
 ||| |||:
 Db 220 AspProLeuGluLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 239
 QY 850 GACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
 ||| |||:
 Db 240 AspLeuGlySerIleLeuCysAspIleLysPheSerCysLeuThrTyLeuAspAlaPhe 259
 QY 910 TGGAGGAGCACTACTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 969
 ||| |||:
 Db 260 TrpArgAspTyrIleLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 279
 QY 970 GCCCTGCGAGAGGCTGTGGCGCGGAGGCTGTTGCGCTGTGCTGCTGCTGCTGCTGCTG 1029
 ||| |||:
 Db 280 SerLeuLysGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 299

Oy		1030 GACATAGAGCTGGCGGCGCCGCTGTG	1059
Dd	300	AspyrTgUleucGIyArgGInlyLeuIeu	309

RESULT 6					
ID	DEDD	RAT	STANDARD;	PRT;	318 AA.
AC	0922K0:				
DT	28-FEB-2003	(Rel. 41,	Last sequence update)		
DT	05-JUL-2004	(Rel. 44,	Last annotation update)		
DE	Death effector domain-containing protein [death effector domain-containing testicular molecule].				
GN	Name=Deed; Synonyms=Delft;				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
RN	NCBI_TaxID=10116;				
RP	[1]				
SC	SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.				
RX	TISSUE=Testis;				
RX	MEDLINE=99049260; PubMed=9932420; DOI=10.1210/en.139.12.4839;				
RA	Leo C.P., Hsu S.Y., McGee B.A., Salanova M., Hsueh A.-J.W.;				
RL	"DBFT, a novel death effector domain-containing molecule predominantly expressed in testicular germ cells.";				
RL	Endocrinology 139:4839-4848(1998).				
CC	-1- FUNCTION: A scaffold protein that directs CASP3 to certain substrates and facilitates their ordered degradation during apoptosis. May also play a role in mediating CASP3 cleavage of KR18. Regulates degradation of intermediate filaments during apoptosis. May play a role in the general transcription machinery in the nucleus and might be an important regulator of the activity of GFP3C3. Inhibits DNA transcription in vitro (By similarity).				
CC	-1- SUBUNIT: Interacts with CASP8, CASP10, KRT9, KRT18, CASP3 and FADD. Homodimerizes and heterodimerizes with DEED2 (By similarity).				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus during C95-mediated apoptosis where it is localized in the nucleoli. Following apoptosis induction, the mono and/or dibiunitation form increases and forms filamentous structures that colocalize with KRT9 and KRT18 intermediate filament network in simple epithelial cells (By similarity).				
CC	-1- TISSUE SPECIFICITY: Widely expressed with highest levels in testis. Within the testis, highly expressed in germ cells but not expressed in Sertoli cells.				
CC	-1- DEVELOPMENTAL STAGE: First detected in 20-day-old animals. Reaches a peak at 30 days.				
CC	-1- PTM: Exists predominantly in a mono- or dibiunitated form.				
CC	-1- SIMILARITY: Contains 1 death effector (DEED) domain.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL;	AF053362;	AAC80287.1;	-	
DR	GO;	GO:0005737;	C:cytoplasm;	ISS.	
DR	GO;	GO:0005730;	C:nucleolus;	ISS.	
DR	GO;	GO:0003677;	F:DNA binding;	ISS.	
DR	GO;	GO:0006917;	P:induction of apoptosis;	ISS.	
DR	GO;	GO:0016481;	P:negative regulation of transcription;	ISS.	
DR	InterPro;	IPRO11029;	DEATH_like.		
DR	InterPro;	IPRO01875;	DED.		
DR	Pfam;	PF01335;	DED; 1.		
DR	SMART;	SMO0031;	DED; 1.		
DR	PROSITE;	PS50168;	DED; 1.		
KW	Apoptosis; DNA-binding; Nuclear protein; Repressor;				
FT	Transcription regulation.				
DOMAIN	25	103	DED.		

[illegible]

RESULT 7

ID	DEDD_MOUSE	STANDARD;	PRT;	318 AA.
AC	O992L3; 07TOH8; Q9R2Z7;			
DT	26-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Death effector domain-containing protein (DEDDProl).			
CN	Name=Deedd;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCL TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.			
RX	MEDLINE=36447559; PubMed=9774341; DOI=10.1093/emboj/17.20.5974;			
RA	Stegh A.H., Schickling O., Ehret A., Scalfitt C., Peterhaensel C.,			
RA	Hofmann T.G., Grumet I., Krammer P.H., Peter M.E.;			
RT	"DEDD, a novel death effector domain containing protein, targeted to			
RT	the nucleolus."			
RL	EMBO J. 17:5974-5986(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Thome M., Teschopp J.;			
RT	"DEDProl, a novel DED-containing protein.";			
RT	Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=FVB/N-3; TISSUE=Colon, and Mammary gland;			
RC	MEDLINE=22388229; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,			
RA	Alecunil S.F., Zeeberg B., Buettow K.H., Scheefel C.F., Bat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatcenco L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stepston M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,			
RA	Brownstein M.J., Udell T.B., Toshynki S., Carlini P., Prange C.S.,			
RA	Raha S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Boek S.A., McEwan P.C., McKernan K.U., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Holyk S.W.,			
RA	Villalon D.K., Muzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Ketteman M., Maden A.C., Rodriguez S., Sanchez A.,			
RA	Whiting M., Maden A., Young A.C., Shechenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodríguez A.C., Grimmood J., Schmütz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Small D.E.,			
RA	Scherer A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
-1-	FUNCTION: A scaffold protein that directs CASPs to certain			
CC	substrates and facilitates their ordered degradation during			
CC	apoptosis. May also play a role in mediating CASP cleavage of			
CC	KRT18. Regulates degradation of intermediate filaments during			
CC	apoptosis. May play a role in the general transcription machinery			
CC	in the nucleus and might be an important regulator of the activity			
CC	of GRF33 (By similarity). Inhibits DNA transcription in vitro.			
-1-	SUBUNIT: Interacts with CAS9, CASP10, KRT8, KRT18, CASP3 and			
CC	similarity).			
-1-	SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus			
CC	during CD95-mediated apoptosis where it is localized in the			
CC	nucleoli. Following apoptosis induction, the mono and/or			
CC	dubulinization form increases and forms filamentous structures			
CC	that colocalize with KRT8 and KRT18 intermediate filament network			
CC	in simple epithelial cells (By similarity).			
-1-	TISSUE SPECIFICITY: Ubiquitously expressed.			
-1-	PTM: Exists predominantly in a mono- or diphosphorylated form.			
-1-	SIMILARITY: Contains 1 death effector (DED) domain.			

[illegible]

```

QY 610 CGCCCACTGCTGTCCTCCAGACGGCGCGGAGAGGGGCCCCAGCCGACCCAGCAGCAG 669
Db 170 ATGTTTThLeuGlySerGlnArgValArgValArgValArgValArgValArgValArgVal 180
QY 670 TCAGAGCCCGCCAGACCTTCTCTGTAAGGCAAGTACCTGTGACATCCGGCTCCGGGTT 729
Db 181 ---LysSerValThrProAspProLysGluValGlnThrCysAspIleArgLeuArgVal 199
QY 730 CGAGCAGACTACTGCGAGCAGATGGCGCCGCTTGAGCAGAGCGCGGTGACATCCCGCGGCC 789
Db 200 ATGAlaGluTyrCysGlnIleGlnThrAlaLeuGlnGlnValAsnValPheSerAsnLysGln 219
QY 790 CAGCGCTCGCCCGGAGCTGACGCTGTTTGGCAGAGCCAGCCGACGCTGCTGGCTCAAG 849
Db 220 AspProLeuGluArgGlnPheGlnAlaGlnPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
QY 850 GACCTGCGCTGCTGTTGTGACATCAAGTTCTGACAGCTCTCTATCTGACGCGCTTC 909
Db 240 AspLeuGlySerIleIleCysAspIleLysPheSerGlnLeuThrTyrLeuAspAlaPhe 259
QY 910 TGGGGCGACTACTGAGTGGCGCCCTGCTGACAGCCCTGCGGGCGGTGTTCTGACTGAG 969
Db 260 TPArgAspTyrIleAsnGlySerLeuLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
QY 970 GCCCTCGCAGAGGCTTGGCGCGGAGGCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029
Db 280 SerLeuLysGlnAlaValGlnAlaIleLysGlnLeuValAsnValAspGlnGln 299
QY 1030 GACTATGAGCTGCGCGCGCGCTGTTG 1059
Db 300 AspTyrGlnLeuGlnLysArgGlnLysLeuLeu 309

RESULT 8
Q8BRM9 PRELIMINARY; PRT; 168 AA.
ID Q8BRM9;
AC Q8BRM9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 10 days neonate cortex cDNA, RIKEN full-length enriched
DE library, clone:A830049M19 product:similar to death effector domain-
DE containing and DNA-binding protein 2, full insert sequence.
DE Name:Dcd2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

```

```

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multichannel capillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume M.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaubuka T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Konda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohatao N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK043908; BAC31699.1; -.
DR GDB; MGI:1914629; Dcd2.
DR GO; GO:0042981; P:regulation of apoptosis; TAS.
KW DNA-binding.
SQ SEQUENCE 168 AA; 17242 MW; F75D71DC489D2425 CRC64;

Alignment Scores:
Pred. No.: 6,63e-20 Length: 168
Score: 530.00 Matches: 125
Percent Similarity: 77.38% Conservative: 5
Best Local Similarity: 74.40% Mismatches: 34
Query Match: 14.99% Indels: 7
DB: Gaps: 3

US-10-030-271-3 (1-1883) x Q8BRM9 (1-168)
QY 473 ATGGCACTCTCAAGCT--CTTCAAGAGCAGAGGCTGCTCCGCTCGGCACT 529
Db 1 MetAlaIleProAlaLeuLeuProArgGlnArgThrValAlaGlyValGlyAspArg 20
QY 530 CAGACGCTTGCCTCAATTTCTGACGAGGCTGATGGAGACAGAGCTCCCGCCAGCACC 589
Db 21 ArgAlaValLeuArgThrLeuLeuArgValSerGlyThrGlnAlaProHisGlnProSer 40
QY 590 GGCAGCGCGGAGTGGCGGCGCGGCGGAGTGTGTGCTCCAGACGCGGCGGAGAG--GGG 646
Db 41 GlySerGlyGlyValGlyValAlaAlaLeuProValValProGlySerGlyGlyGlnAla 60
QY 647 CCCAGCGCGCACCAGCAGCAGCTAGAGCCC-----GCCAGACTTCTCTGAGGC 699
Db 61 LeuGlnProArgSerSerThrArgSerThrArgSerThrAlaAlaProSerSerGlyGly 80
QY 700 AAAGTCACTGTGACATCCGGCTCCGGCTTCCAGACAGACTATTCGAGAGATGGCCAGGC 759
Db 81 LysAlaThrCysAspIleArgLeuArgValAlaGlnAlaGluTyrCysGlnIleGlyProAla 100
QY 760 TTGAGCAGAGGCGGTGACATCCCGCGCGGCGGCGGCGGCTGCGGCGGAGCTGAGCTGTT 819
Db 101 LeuGlnGlnGlyVal***SerArgThrProGlnAlaLeuAlaArgGlnLeuAspValPhe 120

```

QY 820 GGGCAGCCACCGAGTGT- GCGCTCAGAGGACCTGGCTCTG- TGGTTTGACATCA 877
 DB 121 GYGINAAThrAlaValAlProAlaValIleGlyProGlyLeuCyStrpCyValAlThrSer 140
 QY 878 AGTTCTCAGAGCTCTCTATTTGACCGCTTCTGGGGGAGTACTTGAAGTGGCCCTGC 937
 DB 141 SerSerInSerCySProlIerPrThrProSerGlyAlaThrAlaValAlaProCys 160
 QY 938 TGCAGCCCTGCGGGCGTGTCC 961
 DB 161 CyArgPrProheGlyAlaCySer 168
 RESULT 9
 Q6GNZ8 PRELIMINARY; PRT; 243 AA.
 ID O6GNZ8;
 AC O6GNZ8;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DE MGC80767 protein.
 GN Name=MGC80767;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_Taxid=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidi T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Klein S., Gerhard D.S.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC073351; AAH73351.1;
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR001875; DED.
 DR Pfam; PF01335; DED.1.
 DR PROSITE; PS50168; DED.1.
 SQ SEQUENCE 243 AA; 2767 MW; 57A2E89CFBD3E0BC CRC64;

Alignment Scores:
 Pred. No.: 5,34e-13 Length: 243
 Score: 398.00 Matches: 92
 Percent Similarity: 53.71% Conservative: 31
 Best Local Similarity: 40.17% Mismatches: 52
 Query Match: 11.26% Indels: 54
 DB: 2 Gaps: 5
 US-10-030-271-3 (1-1883) x Q6GNZ8 (1-243)
 QY 157 TGGAGAGAGATAGTACCTGAGTACTTACCGGATGCTGCTTACCGGTATTTGAG 216
 DB 13 TrpGluGluSpGluCysLeuGluGlyTrpGlyMetLeuSerLeuHisArgMetPheAsp 32
 QY 217 GTGGTGGGGGGCAACTGACCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 276
 DB 33 ValValGlySerGlnLeuThrGlnAsnAspIleAspAlaLeuSerPheLeuLeuHisGlu 52
 QY 277 GCT----- 279
 DB 53 ThrHisProPheThrHisProLeuAspProGlnLeuThrAlaGluGluAlaGly 72
 QY 280 -----CCTGGCGCCGCC----- 291
 DB 73 GluAlaMetProAsnSerAlaLeuLeuSerAlaTrpGlnArgTrpAsnArgGlySerArg 92
 QY 292 -----GGAGGCTTAGCCCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330
 DB 93 ThrLeuAsnAsnLeuAspSerProAspLeuLeuArgProGlyThrGlyLeuLeu 112
 QY 331 CTGAGAGCTGAGGCGCGCGGCGGAGTGGCGGAGCAAGCACTGGCGGCTGGCGGCAATC 390
 DB 113 LeuGluLeuGluArgArgGlyGlyCysAspGluSerAsnThrHisLeuLeuGlnLeu 132
 QY 391 CTGGCGCTGTGGCGCGGCGGAGCTGCTGCTGCGGAGTGGCGGCGGCGGCGGCGGCGG 450
 DB 133 LeuArgValLeuThrArgHisAspLeuLeuProGlyValThrThrGlyArgProArgAla 152
 QY 451 GTGGCTCCAGAGGCTTACTTACTGAGCACTGC-----AGCTTTGAAGAGGAGAGAG 504
 DB 153 ValSerProGluArgGlyTrpGlyTrpGlyProSerIleLeuAspSerAspGlnGlnMetAsp 172
 QY 505 GGTAGTGGCGGCTGCGGCGGAGTCAAGCAAGTTCGAAATTCAGCAAGGTCAGTGG 564
 DB 173 -----ArgCysLeuAsnProAsnProAlaSerThrArgGluGluAsnTrp 187
 QY 565 GAGACAGGCTCCCCCAACCAAGCGGAGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGG 624
 DB 188 GluThrGlySerAsnSerSerGlyArgGlyGlyGlyThrGlnGlyMetGlyHisCys 207
 QY 625 GCAGAGCGGCGGCGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 684
 DB 208 ProLysProLysArgAsnLysValGlyAsnThrProSerGln-----Asn 222
 QY 685 CCTTCTCTGAGAGCAAGTACCTGT 711
 DB 223 ProAsnAsnGlnSerLysValThrCys 231
 RESULT 10
 Q6DHY2 PRELIMINARY; PRT; 244 AA.
 ID O6DHY2;
 AC O6DHY2;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DE Dcd1 protein.
 GN Name=dcd1;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OC NCBI_Taxid=7955;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Prange C.,
 RA Raha S.S., Lotteliello N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Richards S., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Murray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettlemen M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schermer A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Whole;
 RA Strausberg R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC075865; AAH75865.1; -;
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:004281; P:regulation of apoptosis; IEA.
 DR InterPro; IPR011029; DEATH_1like.
 DR InterPro; IPR001875; DED.
 DR Pfam; PF01335; DED; 1.
 DR PROSITE; PS01668; DED; 1.
 SQ SEQUENCE 244 AA; 28051 MW; 043225AE5D79527 CRC64;

Alignment Scores:
 Pred. No.: 1,84e-10 Length: 244
 Score: 349.50 Matches: 88
 Percent Similarity: 46.15% Conservative: 20
 Best Local Similarity: 37.61% Mismatches: 69
 Query Match: 9.89% Indels: 57
 Gaps: 5

US-10-030-271-3 (1-1883) x QSDHV2 (1-244)

QY 157 TGGGAGAGAGATGAGTGGCTGACTACGCGAGTGTGCTTACCGGTATGTTGAG 216
 DB 13 TTPGGLGTLTGTLGTCYLSERISERTYRGTGTLThrLeuSerIleuHISgluIlePheglu 32
 QY 217 GGTGGTGGGCGGCACTGACCGGAGTGGAGTCTGCTTGTCTTGTGATGAG 276
 DB 33 IIEVAGISerGlnLeuThrGluThrAspValGluValLeuSerPheLeuLeuAspGlu 52
 QY 277 GCT-----CCTGGC 285
 DB 53 ThrTy-ProGlyLysHisProLeuAspProGluGlyTyrThrGluAspLeuProProGly 72
 QY 286 GCCGCCCGA----- 294
 DB 73 ProAspGlySerProGluAlaAsnThrProCyProArgLeuLeuLysSerTyrGlnArg 92
 QY 295 -----GGCTTGGCCGGCGCGCGACGACTCTGCGGCACTGGCGCGCAAGCGG 321
 DB 93 MetGlnProGlnLysGluGlyCysSerIleAlaSerArgHisArgProLysSerGlyVal 112
 QY 322 GAGCTCTGCTGAGTGGAGTGGAGCCCGCGGAGTGGCGGAGCAACCTGGCGTGTG 381
 DB 113 GluLeuLeuLeuGluLeuGluArgGlyTyrLeuSerAspAlaSerLeuArgProLeu 132
 QY 382 GGGCAACTCTGCGGTGCTGGCGCGCGCGACGACTCTGCGGCACTGGCGCGCAAGCGG 441

DB 133 LeuGlnLeuLeuArgIleLeuThrArgHisAspValLeuProPheValSerIleLys 152
 QY 442 CCGCGGCGAGTGTCTCCAGAACCTATAGCTATGCGACCTCCAGCTTTCAAGAGACA 501
 DB 153 ArgArgThrValSerProGluArgGlnLysIleAspTyrProGluValAspPheArgGln 172
 QY 502 GAGGGTACTGCCCGTCGCCGCTGGCGAGTCAACAGCTTTCGCAATTTCTAGCAGAGGTG 561
 DB 173 AspArgGluValGlySerAsnThrAsnIleProSerPheGluAsnThrGlnAspHis 192
 QY 562 TGGGAGACAGAGCTCC-----CCCCAACCAAGCGGAGCGGAGTGGGAGC 609
 DB 193 TTPAGGAlaGlySerGlySerSerMetThrSerAlaSerSerAsnArgArgLysArg 212
 QY 610 CGGCGGAGTGTGTGTCGCGAGCGCGGCGGAGAGGGGCCCGCGGAGCCGACCGACGAGC 669
 DB 213 ArgGlyHisHisIleTyrSerArgLysSerArg-----GlyProProGluIleGlnProGln 230
 QY 670 TGAAGCCCGCGACGACTTCTCTGAAGGCAAGTACTGT 711
 DB 231 SerThrPro-----AsnLysValThrCys 238

RESULT 11

ID Q6C708 PRELIMINARY; PRT; 659 AA.
 AC Q6C708;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Similar to tr|Q95YC9 Sus scrofa Basic proline-rich protein.
 GN ORFNames=VALI026191g;
 OS Yarrowia lipolytica CL1899.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OX NCBI_TaxID=284591;
 RX [1]
 RN RP SEQUENCE FROM N.A.
 RC STRAIN=CL1899;
 RG Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
 RA Coffard N., Frangoul L., Aigle M., Anthonard V., Babout A., Barbe V.,
 RA Barnay S., Blanchin S., Beckrich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Catolico L., Contandriou F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantreuve F., Hémeguin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrist A., Kozul R., Lemaire M., Leur I., Ma L., Muller H.,
 RA Nicard J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenn S., Potier S., Richard G.F., Strub M.L., Suleau A.,
 RA Swemene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zentou-Meyer M., Zivanovic I., Bolotin-Pukhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weisenbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts.";
 RL Nature 430:35-44(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CL1899;
 RA Genoscope;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR382130; CAG81510.1; -;
 DR InterPro; IPR003124; WH2.
 DR Pfam; PF02205; WH2; 1.
 SQ SEQUENCE 659 AA; 62466 MW; 2C55088BC9EF0908 CRC64;

Alignment Scores:
 Pred. No.: 1.21e-06 Length: 659
 Score: 276.50 Matches: 176
 Percent Similarity: 33.84% Conservative: 48
 Best Local Similarity: 26.59% Mismatches: 209
 Query Match: 8.06% Indels: 231
 Gaps: 34

US-10-030-271-3 (1-1883) x Q6C7Q8 (1-659)

[illegible]

D	b		364	erSetIIeAlaSerSeVal..SerThrProSerThrProProProAlaLProProAlaPro	383
O	y		720	CCGATGTTCACAGCTCACTTGGCTTCACAGAGAAGCTCGCGGGCTGACTGCTGCTG	661
D	b		384	ProProAlaProGly-----	388
O	y		660	GAGTGCGGCTGGGGGCCCTTCGCCGGCGCCTGCGACACCACACTGGCGCGGCCGACT	601
D	b		389	-----GlyAlaAlaProProIleProGlySerAlaAlaProProAlaProPro	404
O	y		600	CCGCGCGCTGCCTGTGGTTGGGGG-----	562
D	b		405	ProAlaProProAlaGlyPheGlyLaProAlaProProSerPheGlyLaProThrPro	424
O	y		561	CTGACCCCTGCTGAGAAATTGGACAACCTGCTGACTGCGACGGCGACGGCACTAACCTTC	502
D	b		425	-----ProProAlaProProAlaProSer	432
O	y		501	TGTCTCTTTGAAGCTGGAAGTGCATAGCTAAGCTTTGGAGACATGCGCGGG	442
D	b		433	AlaPro-----ProAla	436
O	y		441	CCGCTTGGCGCGACAGTGCGGCAGCAGGTCG-----	397
D	b		437	Pro--ProAlaProProAlaProProProSerGluProProSerThrProAlaGlyProAlaM	456
O	y		396	GCGCAGAGTT-----GCCCGACGACCGCAGGTTGCTCTCGCGCACCTGCC	349
D	b		456	eLPheGlyAlaLProMetProLySerProAlaAlaL-----SerP	470
O	y		348	GCGGCGCTCCAGCTCCAGCAGGAGCTTAGCGCGCTCGCGGCCCGGGCTTAAGCTCCGGC	289
D	b		470	roGLyAlaLProProProProProPro--GlyAlaAlaAlaLProGlyLeuAla-----	486
O	y		288	GAGCGCAGAGAGCTCATCAGACAGAAAGCCGAGACTCCAGCTCGCACTCGTGACTTG	229
D	b		487	-----ProProAlaProProAlaGlnProProSerProGlyArgProSerGlyAlaP	504
O	y		228	CCCCGCCACCACT-----CGAACATAACGGTGAAGCAGAC	193
D	b		504	roProProProProProGlyProProProAlaProProThrAspGlnHenHISerMetIleLeu	524
O	y		192	CATCCCGTAGTAGTCAGGCACTACT-----CTCTCTCCACACCGGGC	148
D	b		524	sPaSPgLYSerSerSerGlySerHisGlyAlaProProProProProProSerAlaProP	544
O	y		147	CG-----GGTCGACCCCGATGAGGCCATTCGC-----	119
D	b		544	roSerAnGLyGLYhiSserHisGlyAlaProProProProProProProAnglyGlyVaLa	564
O	y		119	-----	119
D	b		564	snArGaRgaRgaSpAlVaLaGlnAgThrSerThrLeuGlySerAnenIleArgThrl	584
O	y		118	-----GGGG	115
D	b		584	euApSThrSerAlaTrhTrIleAlaProArgAlaValSerThrProValGlySerSers	604
O	y		114	AGGAGAGGGGGAACAGCTCAGAACCCGGCCAAATTCTGGCGCTCCATGACTCTCG	59
D	b		604	erGLyGLYGLYGLYGLYSerYsProProGlnIleLeuIleAspaSPSerArgTTP	622
RESULT 12					
O41935					
ID	O41935	PRELIMINARY;	PRT:	585 AA.	
AC	O41935;				
DT	01-JAN-1998	(TREMBLrel. 05, Created)			
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)			
DE	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DR	Hypothetical protein GAMMAHV.M6.				
GN	Name=GAMMAHV.M6;				
OS	Murid herpesvirus 4 (MuHV-4) (Murine gammaherpesvirus 68).				
OC	Virtues; dsDNA viruses, no RNA stage; Herpesviridae;				

OC Gamaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=33708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WUMS;
 RX MEDLINE=97366649; PubMed=9223479;
 RA Virgin H.W. IV, Latreille P., Mamajley P., Hallsworth K., Weck K.E.,
 RT "Complete sequence and genomic analysis of murine gammaherpesvirus
 68."
 RL J. Virol. 71:5894-5904 (1997).
 DR EMBL: U97553; AAB66392.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 585 AA; 60160 MM; 85610AA80C34827D CRC64;

Alignment Scores:

Pred. No.:	2,2e-06	Length:	585
Score:	271.50	Matches:	177
Percent Similarity:	30.80%	Conservative:	38
Best Local Similarity:	25.36%	Mismatches:	223
Query Match:	7.91%	Indels:	261
DB:	2	Gaps:	38

US-10-030-271-3 (1-1883) x 041935 (1-585)

QY 1818 CCAGAGCTTTGATTCCTTCCAGAGAGTGCTGTTTACTGTTTGGCCCATGTG 1759
 DB 13 ProArgProLeuHisIle-PheProLeuPheTrpArgSerIleLeuValGlyProThrCy 32
 QY 1758 CAACACTAGAGCTTGATGATGCTCCATACACTCCCATGTGACATCCAGGTACAT 1699
 DB 32 s-----ValAlaIlePro-----GlyGlyHisIle 40
 QY 1698 CCAAGTGAAGTCTAT-----GTTCACTCTGACAGACAGAGGGGAA 1657
 DB 40 uIysValIleuLeuPheAspValPheTyrSerAspIleIysAlaPheThrSer----- 57
 QY 1656 GGTGAGAGAGTCAAGTCTGTAAATTGAAGTGGGAGGCTGCTGCTGAGATGT 1597
 DB 58 -----LeuSerGlySerPh 62
 QY 1596 GTGGGCAAGTGACAGAGCCCATGTGCAC----- 1566
 DB 62 eTrp-----PheCysSerSerThrCysAlaAspArgPheGlyPhePr 76
 QY 1565 -----CCAGCTCCAGTGCCTCATGATGTGCTGA 1537
 DB 76 oSerAlaGlaGlyAlaArgIleProAspLeuProGlyProLeuProSer----- 92
 QY 1536 AGCCAGGTTGGTTGTCAATTAGAGTCTGAGAGAGAAAGCCCTTAGGCCAAAGAA 1477
 DB 93 -----TrpGlyProAsp----- 96
 QY 1476 GTGTGTGCTGGAGAGAGGCGCAGAGAGAGCCACAGGAAGATCTACAGAGGGGCACTGT 1417
 DB 97 -----ProArgPro-----ProArgProProProGluLeuGlyProGlySer-- 110
 QY 1416 GAGCATGAGAGAGTGTGCACTGTGACAGTGCAGACAGACCCCAAGATCAGAGATCAAT 1357
 DB 111 -----ProThrSerProAlaProSe 117
 QY 1356 GTGCAGAGGGGCTTTGTGAGTGTAGTGTGAGCCCTCCCTTC-----GAGAT 1306
 DB 117 rArgAlaGlyAlaArgIleProAspLeuProGlyProLeuProSerTrpGlyProAspPr 137
 QY 1305 ACAAGGCCAGAGCCCGCTCCGAGAGCTAAGGGGAGACACTGGGGGTGAGAGAGTCTG 1246
 DB 137 oArgProProArgProProGlu-----LeuGlyProGlySer----- 150
 QY 1245 GGAAGAACTGAGTGAAGATGTTAGGCTCAGAGCCACATCTGTGGAGGGGCT 1186
 DB 151 -ProThrSerProAlaProSerArgAlaGlyAlaArgIleProAspLeuProGlyProLe 170

QY 1185 G---TCAAGGGGTAGAGATGTCCTCTCCAGAGAAAGTGAGCCCGAGACTTGAGGT 1129
 DB 170 uProSerTrpGlyProAspProArgProProArgPro-----ProProGluLeuGlyPr 188
 QY 1128 GGGATCA----- 1122
 DB 188 oGlySerProThrSerProAlaProSerArgAlaGlyAlaArgIleProAspLeuProGl 208
 QY 1121 -ATCTGCCAGTCTTGATCAGAGAGCCCTGTGCGGGCGCCGCCCTTCTCTCCAT 1063
 DB 208 yProLeuPro-SerTrpGlyProAspPro-----ArgProProArgProProProG 225
 QY 1062 CAGCAACAGAGCGCGCCGCGCCAGCTCATATGACCTTCATCAACTGACACAGAGGG 1003
 DB 225 IuLeu-----GlyProGlySerProThrSer--ProAlaProSerArgAlaGlyAlaA 242
 QY 1002 AACAGCT-----CCGAGCCAGAGCCCTCTGACAGGGGCTCAGTCAGGAACAGCGCCG 949
 DB 242 rGlyLeProAspLeuProGlyProLeuProSerTrpGlyProAsp-----ProArgPro 260
 QY 948 CAGGCTGTGACAGAGGCGCCACTCAGTATGTGCGCCAGAGAGGCGTCCAGATAGAGAG 889
 DB 260 roArgProProProGluLeuGlyProGlySerProThrSerProAlaProSerArgAlaG 280
 QY 888 CTCTGAGAACTTGATGTCAAAACACAGAGCCCGAGTCCCTTGAGCGACACTGCGGT 829
 DB 280 lValAlaArgIle-----ProAspLeuProGlyProLeuProSer-----T 293
 QY 828 GGCCTGCCAAACAGTCCAGTCCGCGCGCCAGCGCGCGCGCGGATGCCA---C 772
 DB 293 rGlyLeProAspProArgProProArgProProProGluLeuGlyProGlySerProThrs 313
 QY 771 GCCCTGCTCCAGAGGTGCGCCATGCTCGAGTACTGTCTGACACCCGAGCGGATGTC 712
 DB 313 eProAlaProSerArgAlaGlyAlaArg-Ile-----ProAspLeu 326
 QY 711 ACAGTCACTTTCCTTCAGAGAAAGTGTGCGGCGCTGTGACTGCTGCTGGGGTGGCGC 652
 DB 327 ProGlyProLeuPro-----Ser 332
 QY 651 TGGGGCGCCCTCTCCGCGCGG-----TGTGGCACACCACT 616
 DB 333 TrpGlyProAspProArgProProArgProProProGluLeuGlyProGlySerProThr 352
 QY 615 GGGCGCGCCCGACT-----CCGCGCTGCGGCTT 586
 DB 353 SerProAlaProSerArgAlaGlyAlaArgIleProAspLeuProGlyProLeuPro-- 371
 QY 585 GGTGGGGGGGAGCCTG-----TCTCC 563
 DB 372 SerTrpGlyProAspProArgProProArgProProProGluLeuGlyProGlySerPro 391
 QY 562 ACTGACCCGTGAGAAATTGACAGAACTGTGACTGCGAGAGGGAGCGGACATACCT 503
 DB 392 ThrSerProAlaProSerArgAlaGlyAlaArgIleProAspLeuProGlyProLeuPro 411
 QY 502 CTGTCTCTTTGAGAGTGTGAGTGCATATAGCTTATAGCTTGTGAGACACTGGCGGCG 443
 DB 412 -----SerTrp----- 413
 QY 442 GCCGCTTGCGCCAGAGTGCAGAGAGTCTGTGCGGCGGACGACAGCGAGAGTGGC 383
 DB 414 -----GlyProAspProArgProProArgProProProGluLeuGlyProGlySer 430
 QY 382 CCAGAGCCGAGAGTGTCTGTGCGGCACTGCGCGCGGCGCTCCAGCTCCACAGAGAGT 323
 DB 431 ProThrSerProAlaProSerArgAlaGlyAlaArg--IleProArgProSerArgPro 449
 QY 322 CTAGCGCGCTGCGGCGCGGCTA---AGCTCCGCGGCGGCGCAGAGAGCTCATCAGCA 266
 DB 450 ProProGluLeuGlyProGlySerProArgPheProArgProSerArgProProProGlu 469
 QY 265 GAAAGGCCAGAGCT----- 251

Db 343 eGlyProAlaAspProProAlaAlaAlaArgLeuProProGlyGlnGluProArgL 363
 QY 561 -----CTGACCCCTGCTGAGATTTCAGAACTGCTTACTGCCGACGCGACGG 513
 Db 363 euProGlnAspLeuAlaAlaAlaGlnArgCysProAlaGlyProProProThrArgSerg 383
 QY 512 CAGCTACCTC----- 502
 Db 383 lYAlaAlaAlaGlnArgThrHlsArgArgProProGlyCysProArgSeraAlaArgAsp 403
 QY 501 -----TGTCCT-----CTTGAAGAGCTGA-----G 480
 Db 403 roGlyCysProArgThrTPArgArgArgSerglYAlaGlnArgGlyHlsProProProG 423
 QY 479 GTGCCATAGCTATAGCGTTCTGGAGACACTGGCGCGCGCTTGGCGCCAGCTGGCGC 420
 Db 423 lYAlaGlyGlnArgPro-SerglYProThrGly-GlyArgProAlaAlaProGlyAlaArg 442
 QY 419 AGCAGGTCTGGCGCGCGCCAGACGCGGAGCTGCCAGACGCGCGAGCTTGTCTCG 360
 Db 442 oGlyThrProAlaAlaProGlyProGlyGlyAlaAlaAlaProSerglYAlaThrPr 462
 QY 359 CCGCACTGCCGCGCGCT---CCAGCTCAGACGAGCTTACGCGCGCGCGCGCG 303
 Db 462 oHlsProGlnArgGlySerglYProAlaAspProProAlaAlaAlaArgLeuProPro-- 481
 QY 302 GCTAAGCTCCGCGCGCGCGAGAGCTCAT-----CCAGCAGAAAG 261
 Db 482 -----GlnArgGlnGluProProGlnGluProGlnAspLeuAlaAlaGlnArg 497
 QY 260 G---CCAGAGAGCTCCAGCTCGCACTGCTGTCAGTTGCC-----CG 225
 Db 497 gCysProAlaGlyProProProThrArgSerglYAlaAlaGlnArgThrHlsArgAr 517
 QY 224 CCACCA-----CCTCGAAGATACGCTGAAC 198
 Db 517 gProProGlyCysProArgSeraAlaArgAsnProGlyCysProArgThrTPArgArgAr 537
 QY 197 GACAGATCCCTAGTAGTCACGACCTCATCTCTCCAGCAGCGGG----- 149
 Db 537 gSerglYAlaGlnArgGly-----HlsProProProGlyAlaGlyGlnArgProse 554
 QY 148 -----CCGGGTGACCCCGATAGCCCATTCGCGG 117
 Db 554 rGlyProThrGlyGlyArgProAlaAlaProGlyAlaProGlyThrProAlaAlaProG 574
 QY 116 GGAGGAGGCGGAGACAGCTCAGAACCCGCGC 86
 Db 574 yProGlyGlyGlyAlaAlaAlaProSerglY 584
 RESULT 14
 ID 0777A3 PRELIMINARY; PRT; 660 AA.
 AC 0777A3;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE BHLPI early reading frame.
 OS Human herpesvirus 4 (Epstein-Barr virus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Lymphocryptovirus.
 NX NCBI_TaxId=10376;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=84270667; PubMed=6087149;
 RA Baer R.J., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
 RA Gibson T.J., Hatfull G.F., Hudson G.S., Satchwell S.C., Seguin C.,
 RA Tuftnell P.S., Barrett B.G.;
 RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
 RL Nature 310:207-211 (1984).
 RN (2)
 RP SEQUENCE FROM N.A.

RC STRAIN=B95-8;
 RX MEDLINE=88283646; PubMed=2840285;
 RA Laux G., Pericardet M., Farrell P.J.;
 RT "A spliced Epstein-Barr virus gene expressed in immortalized
 RT lymphocytes is created by circularization of the linear viral
 RT genome";
 RL EMBO J. 7:769-774 (1988).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=82014887; PubMed=6269068;
 RA Arrand J.R., Rymo L., Walsh J.E., Bjorck E., Lindahl T., Griffin B.E.;
 RT "Molecular cloning of the complete Epstein-Barr virus genome as a set
 RT of overlapping restriction endonuclease fragments";
 RL Nucleic Acids Res. 9:2999-3014 (1981).
 RN (4)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=82059504; PubMed=7301588;
 RA Kozak M.;
 RT "Possible role of flanking nucleotides in recognition of the AUG
 RT initiator codon by eukaryotic ribosomes";
 RL Nucleic Acids Res. 9:5233-5252 (1981).
 RN (5)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=83109311; PubMed=6296170;
 RA Deininger P.L., Bankier A., Farrell P., Baer R., Barrett B.;
 RT "Sequence analysis and in vitro transcription of portions of the
 RT Epstein-Barr virus genome";
 RL J. Cell. Biochem. 19:267-274 (1982).
 RN (6)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=83169725; PubMed=6300857;
 RA Farrell P.J., Deininger P.L., Bankier A., Barrett B.;
 RT "Homologous upstream sequences near Epstein-Barr virus promoter";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1565-1569 (1983).
 RN (7)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=85035713; PubMed=6092825;
 RA Bankier A.T., Deininger P.L., Farrell P.J., Barrett B.G.;
 RT "Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
 RT Epstein-Barr virus";
 RL Mol. Biol. Med. 1:21-45 (1983).
 RN (8)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=85060424;
 RA Seguin C., Farrell P.J., Barrett B.G.;
 RT "DNA sequence and transcription of the BamHI fragment B region of B95-
 RT 8 Epstein-Barr virus";
 RL Mol. Biol. Med. 1:369-392 (1983).
 RN (9)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=83294686; PubMed=6310141;
 RA Jeang K.T., Hayward S.D.;
 RT "Organization of the Epstein-Barr virus DNA molecule. III. Location of
 RT the P3HR-1 deletion junction and characterization of the NotI repeat
 RT units that form part of the template for an abundant 12-O-
 RT tetradecanoylphorbol-13-acetate-induced mRNA transcript";
 RL J. Virol. 48:135-148 (1983).
 RN (10)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8;
 RX MEDLINE=85060428; PubMed=6094955;
 RA Bankier A.T., Deininger P.L., Satchwell S.C., Baer R., Farrell P.J.,
 RA Barrett B.G.;
 RT "DNA sequence analysis of the EcoRI Dhet fragment of B95-8 Epstein-
 RT Barr virus containing the terminal repeat sequences";
 RL Mol. Biol. Med. 1:425-445 (1983).

RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=20331131; PubMed=10872327;
 RA Farrell P.J., Bankier A., Seguin C.,
 RT "Latent and lytic cycle promoters of Epstein-Barr virus";
 RL EMO J. 2:1331-1338(1983).
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=84207939; PubMed=6327290;
 RA Jones M.D., Foster L., Shedy T., Griffin B.E.;
 RT "The EB virus genome in Daudi Burkitt's lymphoma cells has a deletion
 RT similar to that observed in a non-transforming strain (P3HR-1) of the
 RT virus";
 RL EMO J. 3:813-821(1984).
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=84236104; PubMed=6203743;
 RA Buggin M., Farrell P.J., Barrell B.G.;
 RT "Transcription and DNA sequence of the BamHI L fragment of B95-8
 RT Epstein-Barr virus";
 RL EMO J. 3:1083-1090(1984).
 RN [14]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=84222045; PubMed=6328526;
 RA Yates J., Warren N., Reisman D., Sugden B.;
 RT "A cis-acting element from the Epstein-Barr viral genome that permits
 RT stable replication of recombinant plasmids in latently infected
 RT cells";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:3806-3810(1984).
 RN [15]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=84247360; PubMed=6330697;
 RA Gibson T., Stockwell P., Ginsburg M., Barrell B.;
 RT "Homology between two EBV early genes and HSV ribonucleotide reductase
 RT and 38k genes";
 RL Nucleic Acids Res. 12:5087-5099(1984).
 RN [16]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=87289053; PubMed=3039467;
 RA Bodescot M., Perricaudet M.;
 RT "Clustered alternative splice sites in Epstein-Barr virus RNAs";
 RL Nucleic Acids Res. 15:5887-5887(1987).
 RN [17]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RX MEDLINE=91021036; PubMed=2171209;
 RA Parker B.D., Bankier A., Sacchewell S., Barrell B., Farrell P.J.;
 RT "Sequence and transcription of Raji Epstein-Barr virus DNA spanning
 RT the B95-8 deletion region";
 RL Virology 179:339-346(1990).
 RN [18]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RA Hatfull G.F., Barrell B.G., Quinn J., McGeoch D.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 RN [19]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B95-8.
 RA Birnie U.K., Aron W., Farrell P.J.;
 RT "Induction of Epstein-Barr virus late promoters on small plasmids in
 RT the EBV late lytic cycle requires ori lyt.";
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ507799; CAD53473.1; -.
 SO SEQUENCE 660 AA; 66244 MW; 86DA1D67A37152A2 CRC64;

Score:	269.00	Matches:	167
Percent Similarity:	31.60%	Conservative:	23
Best Local Similarity:	26.38%	Mismatches:	231
Query Match:	7.84%	Indels:	202
DB:	2	Gaps:	32
US-10-030-271-3 (1-1883) x Q777A3 (1-660)			
QY	1577	CCCATGTGCACCCAGCTCCAGTGCACATGTGCTGAAAGCCAGATTGGTTGTGGT	1518
DB	21	ProProCybLeuProGlyAlaProAspGlnGlnThrArgLeuProProGlyTyrGly	40
QY	1517	CAATTAGAGTCTGTGAGAG-----GAAACCCCTTAGCCCAAGAAAGTGT---	1473
DB	41	GlnArgThrAlaProThrGlnValGlyLeuAlaAspAlaAspProAspGlyLeuGln	60
QY	1472	-----GNGCTGGAGAGAGGCGCAGAGGCGCAGAGGCGGAGGAAGATTCAGAGGGGCGAGT	1419
DB	61	AspGlnAlaAspGlyAlaArgProGlyGlyGlyAsnArgValGlyAlaGlyValGly---	79
QY	1418	GTGAGCATGAGCAGAGTGTGTGCACCTGTGACAGTGCACAGGCCCAAGATTCAGATATCA	1359
DB	80	-----ArgPro-----	81
QY	1358	ATGTGCAGAGGGGCGCTTTGGTAGTGTGTAGTGTGCTGCCCTCTTGATATCAGGCC	1299
DB	82	-----GlyThrProAlaProSerArgGlnSerArg	91
QY	1298	CAGCCCCCGCTCCGGAG-----GTTAAGGGGACACATCGGG-----GTAGAGAGAGTCTGG	1245
DB	92	ArgThrGlyProAlaGlnGlnAlaAspHisValHisSerAsnProThrGlyGlyCysSer	111
QY	1244	GAA-----GGAACATCAGCTGGAATGGTTTAGGCTCAGAGCCAC	1203
DB	112	AspProGlnArgSerProArgThrArgGlnAlaGlyTyrAlaLeuGly-----	127
QY	1202	ATCTGTGGAGGGGCTGTCAAGGGGTAGAGANTGTGTCTCTCCAGAGAGAGGTGGCC	1143
DB	128	-----GlnGly-----SerAlaGlyLeuGlySerArgGlyProArgProHisProAla	143
QY	1142	GGAACATTGAGAGTGG-----GATCATCTGTGC---CAGTCTGGATCAGGAAGCC	1095
DB	144	PheGlnValGln-TripSerAlaArgAsnProGlyCysProArgThrTyrArgArgArg--	162
QY	1094	TCTGTGGGGCGCGCCCTCTCTCTCTCCATCAGCAAGAGCGCGCGCGCAGCTCA	1035
DB	163	-----SerGlyAlaGlnArgGlyHisProProPro-----GlyAlaGlyGlnArgPr	178
QY	1034	TAGTCAGCTCATTCACACATGCACGAGCAGAGGGAACAGCTCCCGGCCACAGCTTCGC	975
DB	178	oSerGlyProThrGlyGlyArgProAlaAlaProGlyAlaProGlyLysProAlaAlaPr	198
QY	974	AGGCGCTCAGTCAGGAACAGC---CCCGCAGGCGCTGCAGAGGGCGCCATCAGGTAG	918
DB	198	oGlyProGlyGlyGlyAlaValProSerGlyAlaThrProHisProGlnArgGlySe	218
QY	917	TGCGCCCAAGAGGGGTCCAGATGAGAGAGCTGTGAGAACTTGAATGTACAAACCAAGAG	858
DB	218	rgLProAlaAspProProAlaAlaAlaArgLeuProProGlnArgGlnGlnProArgLe	238
QY	857	CCCAAGTCCCTTGAAGCGCAGACATCGAGTGGCTGCCCAACAGTCCAGTGGCGCGC	798
DB	238	wProGlnAspLeuAlaAlaGlnArgCysProAlaGlyProProProThrArgSerGly	258
QY	797	AGGCGCTGGGGCGCGCGGATGCCAGCTCT---GCTTCAGAGCTGCCCATGTCTCG	744
DB	258	yAlaAlaAlaGlnArgThrHisArgArgProProGlyCysProArgSerAlaArg----	276
QY	743	CAGTACTCTGTGCAAACCGGAGCCGGAATGTCAAGTCACTTTGCTTCAGAGGAAGGT	684
DB	277	-----AsnProGlyCysProArgThr-----	283
QY	683	CTGCGGGGCTCTGACTGTCT-----GGGGTGGCGGTGGGGC	645

[illegible]

OC Bacteria; Actinobacteria; Actinobacteriales;
 OC Streptomycineae; Streptomyces.
 OC NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NA-4680;
 RX MEDLINE=11477403; PubMed=11572948; DOI=10.1073/pnas.211433398;
 RA Omura S., Ikeda H., Ishikawa T., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hatori M.;
 RT "genome sequence of an industrial microorganism Streptomyces
 RT avermiltis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NA-4680;
 RX MEDLINE=12670836; PubMed=12692562;
 RA Ikeda H., Ishikawa T., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hatori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermiltis.";
 RL Nac. Biotechnol. 21:526-531(2003).
 DR EMBL; AP005035; BAC71268.1; -.
 DR GO; GO:0005737; C:cycloplasm; IEA.
 DR GO; GO:0006817; P:phosphate transport; IEA.
 DR InterPro; IPR008160; Collagen.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 903 AA; 87231 MW; 11AE6427936489DF CMC6;

Alignment Scores:					
Pred. No.:	6.49e-06	Length:	903		
Score:	262.50	Matches:	182		
Percent Similarity:	31.75%	Conservative:	38		
Best Local Similarity:	26.26%	Mismatches:	261		
Query Match:	7.43%	Indels:	212		
DB:	2	Gaps:	37		

US-10-030-271-3 (1-1883) × Q82HF3 (1-903)

OY	33	CCHAAACAGTGGCTGCGGCSTGCCGCCAGAGATCATCGGACGGCCAGAACATCTGCGCGGT	92
Db	172	ProArgThrValAlaProAlaLeuAspAspAlaValProAlaIla-----ProGln	189
OY	93	TTCGACTTGTTCCGCGCTCCTCCGCCGGAATGCGCTATCCGAGTGAACCCCGGCC	152
Db	190	GlyProValThrProProProGlyPro-----GlyAlaProAlaGvalProserGlyPro	207
OY	153	GTGCTGGAGAGAGATGAGTCCCTGAGCTACTACCGGATGCTTGCGCTTCAACG----	206
Db	208	SerAlaAlaGlnGly-----ProGlyAlaProGluGlnAlaAlaProSerAlaProSer	225
OY	207	-----TAGTTGAGAGCTGCTGGGGGGCAACTGACCGAGTGGAGCTGG-----	251
Db	226	TyrGlyTyProGlnGlyProGlyAlaAlaProSerAlaProSerTYrGlyTyTyProGln	245
OY	252	GCTCTGGCGCTTTCTGCTGAGTGAAGAGCTCTGGCGCGCCGCGAAGCTTAAGCCCCGGC--	308
Db	246	GlyProGlyAlaValAlaProGly-----ThreProProGlyAla	255
OY	309	CGGACGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGCGGACATGCGCGGAGC--	367
Db	259	ProGlnAlaProGlyAlaProHisAlaProGlyAlaProGlnAlaProThryGlySerPro	278
OY	366	-----CAACTGGCGGCGTGGG	383
Db	279	ValProGlyArgProLeuAlaProAsnAlaGlySerIleAlaSerAlaAlaThrSerLys	296
OY	384	GCAACTCT-----GCGGTGCTGGCGCGCACGACCTGTGCGCGCACTGGCGG	433
Db	299	AlaAlaProProArgGlyAlaAlaArgAlaGlyAlaProAlaGProProProGlyAlaProGlyVal	318
OY	435	CAAGCGCGCGCGCAGTGTCTCCAGAAGCSTAATGATGAGCACTTCAGCTTTCAA	494

```

Db      319  |||||  ...  |||||  |||||  |||||  |||||
      495  GAGGACAGAGGATAGTCGCGCGTCGCGGACGTCAAGCAATTCTGCAATTCTGACGA 554
Db      334  ----- 334
      555  GGGTCAGTCGAGAGCAGAGCTCCCCCAACAGCGGACGCGGAGTCGGGCGCGCC 614
      335  -----SerGlyProGlyAlaProGlyThrProAlaGly-----GlyThrValP 349
Qy      615  CAGTG--GTGGTCCAGACGCGCGGAGAGGGGCCAG-----CCGCACC 659
Db      349  roThrgInLeuValSerGlnLeuGlyProAlaPrglyProGlnGlyLeuAlaGlyProAlaP 369
Qy      660  CCAGACAGTCAGACGCGCGGACGCTTCTCTGAAAGGCAAGTACCTGTGACATCCG 719
Db      369  roProSerAlaProGlyThrProAlaPPro-----GlyGlyThrProP 384
Qy      720  GCTCCGCGTTTCAGACAGTACT-----GCGACATGCGCCAGCTTGGACAGG 770
Db      384  roGlyGlyValHisHisAlaAlaAlaThrMetPheAlaAaPProGlyArgPro-GlyAlaGly 403
Qy      771  CGTGGCATCCCGCGCGCCCAAGCGCTGGCGCGGACGCTGACGTGTTGGGACAGGCCAC 830
Db      404  ProGlyAlaPProGlnProGlyPProGlyPProGlyAlaPProGlyArgProGlyValProAla 423
Qy      831  CGCAGTCCTCGCTCAAGAGGACCTGGGCTCTGTGTCATCAAGTTCTCAGAGCT 890
Db      424  ProProGlyPProProGlyAlaPProGlyGlnProGly----- 435
Qy      891  CTCCTATCTGAGACGCTTCTGGGCGACTACTAGTGGCGCCCTGTGAGCGCTGCG 950
Db      436  -----AlaProGlyAlaProGlyAla 442
Qy      951  GGGCGTCTTCTGACTGAGGCGCTGCGAGAGC-----TGT 986
Db      443  ProAlaThrProGlyGlyThrProProGlyGlyValHisHisAlaAlaThrMetLeuAla 462
Qy      987  GGGCGCGGAGGCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1046
Db      463  GlyProSerValGlyGlyPro-----GlyAlaPro 472
Qy      1047  GCGCGCGCTGTTCGTGATGAGAGGAGGAGGCGCGCGCGCGACAGAGGCTCTGTATC 1106
Db      473  LeuProProProGlyAlaProGlyAlaProGlyMetProProGly--AlaProGlyAlaP 492
Qy      1107  CAGGACTGGACAGA-----TTGATCCCACTC 1133
Db      492  roGlyAlaPProGlyGlyProGlyGlyValHisHisAlaGluThrMetLeuAlaGlyProP 512
Qy      1134  CAACTCTCCGGGCGCA-----CCTTCTCTCGGGA-----GACGACCATCTCAACCCCT 1181
Db      512  ro-ValGlyGlyProGlyLeuProProPro-GlyAlaPProGlyAlaPProGlyMetProAl 531
Qy      1182  TGACAGCGCCCTCC-----CACAGATGTGG-----GCTGTAG 1214
Db      531  aGlyThrProProGlyAlaPProGlnGlyTTPProProProGlnProGlyGlnAlaMetAl 551
Qy      1215  GCCTAAACCAATTTCAGCTGAG-----TTTCTTCCAGACTCTCTCC 1256
Db      551  aProGlnProMetProGlyGlnProMetProProGlnProMetProGlyGlnGlnProPr 571
Qy      1257  TACCCCAAGGTGTCCTTACCC-----TCCGAGAGCGGGGCTGGGCTGTATC-- 1308
Db      571  oAlaTyGlyTyPProProThrGlnProThrGlnProThrValGlyProGlyGlnAlaVal 591
Qy      1309  -----TCAGAAAGGAGGAGGACAGACTACACACTCAACCAAGGC 1346
Db      591  uArgTyArgAlaGlnaPrglySerGlnGlnGlnLeuIleArgArgSerAlaPProGlyTh 611
Qy      1347  CCCCCTGACATATGATCTTGATCTTGGGCTGTCTGACATGTCAAGGTGACACACTC 1406

```

```

Db      611  rPro---HisProGluTrpGlnIleLeuHisGlnLeuValGlyAlaMetAsnValProProG 630
Qy      1407  GCTCAGTCTCACACTGCCCCCTGCTGAGATCTTCCCTGGGCTCTGCGCTGCTGCTGCC 1466
Db      630  nGlnValLeuGlu-----LeuHisThrGlnLeuGlnSerCyGlu-----LeuPr 645
Qy      1467  CAGCACACACTTTTGGC-----CTAAGGCTTCTCTCTAGAGACTTGAATTGACCACA 1523
Db      645  oGlyAlaTyCyAlaArgMetIleArgGlnSerTrp-----ProG 659
Qy      1524  ACCAAGCTGGGCTTCAAGCCATCAAGTGGGCACTGAGCTGGG----- 1566
Db      659  nAlaArgIleThrSerIleAlaThrTyGlyThrAspHisAlaSerArgGlnGlnIyme 679
Qy      1567  -----GTCCACATGCGGCGCTGCTCACTTGGCCACACATCTCCAGACGCA 1613
Db      679  rGlnGlnLeuLeuAlaHisGlnGlyGlnLeuHisGlnValAlaAspGlyProAlaArgPr 699
Qy      1614  GGGCCCTGCGCAGCTTCAATTTCAGACCTGACTGTCTTCACTTCCCTCCCTGCT----- 1668
Db      699  oGlyProValArg-----SerProLeuProProValProProAlaGlnPr 714
Qy      1669  -----GTCCAGACTGACATAGACTTGCACCTTGGATGTCACCTGG 1709
Db      714  oIleProProGluAlaIleAlaGlnGlu-----LeuAlaAlaAlaPheGlyProG 731
Qy      1710  AGTGTCAATGAGAGTGTATGACAGATCATCAAGGCTCTGTGACATGAGGCGCC 1769
Db      731  yValIlePheHisPheAspGlnAlaAlaValaSerArgGlnGlyVal----- 745
Qy      1770  AAAACAGTAAACAGCACCTTCTT 1794
Db      746  -ProProIleValAlaHisThrLeu 753

```

Search completed: March 22, 2005, 01:51:41
 Job time : 347.434 sec

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 01:00:10 ; Search time 67.1055 Seconds
(without alignments)
5399.743 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 3535
Sequence: 1 agcgcgataatagagaag.....ctgaaagcgataggtg99 1883

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p_model -DEV=xdp
-Q=cgnt2/USFTO_epool_b/US10030271/runat_21032005_153458_26480/app_query.faeba_1.3150
-DB=pir_79 -QMT=faetan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=SI0030271.CCEN_1.1.151.@runat_21032005_153458_26480 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOBERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
C 1	269	7.8	660	1 QOBE3	BHLFI protein - hu
C 2	259	7.3	660	1 QOBE3	BHLFI protein - hu
C 3	258.5	7.5	1460	1 EDBE1F	immediate-early pr
C 4	253.5	7.4	1466	1 CGHUTL	collagen alpha 1(I
C 5	245	7.1	1446	1 A45344	immediate-early pr
C 6	237.5	6.9	1585	2 T31611	hypothetical prote
C 7	235	6.6	1042	1 CGCHIS	collagen alpha 1(I
C 8	235	6.6	1414	1 I23809	collagen alpha 1(I
C 9	230.5	6.7	886	2 IS0694	collagen alpha 1(I
C 10	230	6.7	1106	2 J00405	hypothetical 119.5
C 11	229	6.7	1464	2 SS9856	collagen alpha 1(I
C 12	228	6.6	779	1 CGBOIS	collagen alpha 1(I
C 13	227.5	6.4	1453	2 S21626	collagen alpha 1(I
C 14	225	6.6	1042	1 CGCHIS	collagen alpha 1(I

C 15	225	6.6	1049	1 CGBO7S	collagen alpha 1(I
C 16	225	6.6	2715	2 T13049	eyelid - fruit fly
C 17	223.5	6.3	1466	1 CGHUTL	collagen alpha 1(I
C 18	223	6.5	416	1 SKXUAG	dermal gland prote
C 19	223	6.3	460	2 T33110	hypothetical prote
C 20	222.5	6.3	1464	1 CGHUIS	collagen alpha 1(I
C 21	220	6.2	779	1 CGBOIS	collagen alpha 1(I
C 22	219.5	6.2	1585	2 T31611	hypothetical prote
C 23	219.5	6.2	2715	2 T13049	eyelid - fruit fly
C 24	218.5	6.4	1464	1 CGHUIS	collagen alpha 1(I
C 25	217.5	6.2	784	2 J00317	hypothetical 82k p
C 26	217.5	6.3	825	1 EDBEXD	immediate-early pr
C 27	217.5	6.3	839	2 F75518	hypothetical prote
C 28	216.5	6.3	775	1 EDBB11	immediate-early pr
C 29	216.5	6.3	1603	2 S23810	collagen alpha 1(X
C 30	215.5	6.3	317	2 A28996	proline-rich prote
C 31	215.5	6.3	488	2 A27353	collagen alpha 1(I
C 32	215	6.3	3164	1 WMBE86	UL36 protein - hum
C 33	214.5	6.3	460	2 T33110	hypothetical prote
C 34	214.5	6.1	674	2 S13301	collagen alpha 1(X
C 35	214.5	6.1	1418	2 T45467	collagen alpha 1(I
C 36	213.5	6.2	1791	2 T02345	hypothetical prote
C 37	213	6.2	924	2 S27923	gene LFP3 protein -
C 38	213	6.2	1453	2 S21626	collagen alpha 1(I
C 39	212.5	6.0	743	1 S23779	collagen alpha 1(V
C 40	212	6.0	744	1 A33426	collagen alpha 1(V
C 41	212	6.0	744	1 S23298	collagen alpha 1(V
C 42	212	6.2	862	2 T46289	hypothetical prote
C 43	212	6.0	1049	1 CGBO7S	collagen alpha 1(I
C 44	212	6.0	1419	2 A41182	collagen alpha 1(I
C 45	212	6.0	1487	2 B41182	collagen alpha 1(I

ALIGNMENTS

RESULT 1
QOBE3
BHLFI protein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 25-Feb-1985 #sequence, revision 25-Feb-1985 #ext_change 09-Jul-2004
C/Accession: A03742
R/Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A>Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A/Reference number: A93065, PMID:85035713, PMID:6092825
A/Accession: A03742
A/Molecule type: DNA
A/Residues: 1-660 <BAN>
R/Baker, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
Nature 310, 207-211, 1984
A>Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A/Reference number: A03794, PMID:84270667, PMID:6087149
A/Contents: annotation; protein coding region
C/Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-52;
C/Superfamily: human herpesvirus 4 BHLFI protein

Alignment Scores:

Pred. No.: 2.08e-08
Score: 269.00
Percent Similarity: 31.60%
Best Local Similarity: 26.38%
Query Match: 7.84%
DB: 1
Gaps: 32

US-10-030-271-3 (1-1883) x QOBE3 (1-660)
QY 1577 CCATGTCACCCAGCTCAGTCCACATGATGTGGCTGAAGCCAGGTGTGGTGTG 1518
DB 21 ProProCySeubProGlyAlaProAapGInGInThArgArgLeuProProGlyTTPGly 40
QY 1517 CAATTAAGAGCTCTGAGAGA-----GAAGCCTTAGGCCAAGAGGTGT--- 1473

```

Db      41 GlnArgThrAlaProThrGlnValGlyLeuAlaAspAlaAlaSerProAspGluLeuGln 60
Qy      1472 -----GTGCTGGAGACAGCCAGGCGAGAGGCCCGAGGAAGATCTCAGAGGGGAGCT 1419
Db      61 AspGlnAlaSerGlyAlaArgProGlyGlyAlaAsnArgValGlyAlaGlyArgGly---- 79
Qy      1418 GTGAGCATGAGCAGGTGTGTGACCTGTGTGACAGTGCAGACGCCCAAGATCAGAGTATCA 1359
Db      80 -----ArgPro----- 81
Qy      1358 ATGTGACAGGGGGGCTTTGGTGTAGTGTGTAGCTGTGCCCCCTCTGTAGATATACAGGCC 1299
Db      82 -----GlyThrProAlaProSerArgGlnSerArg 91
Qy      1298 CAGCCCCCGCTCCGAG---GCTAAGGGGGCAGACCTGGG---GTAGAGGAGTCTGG 1245
Db      92 ArgThrGlyProAlaGlnGlnAlaAspHisValHisSerAsnProThrGlyGlyCysSer 111
Qy      1244 GAA-----GGAACCTCAGCTGGAAATGTTTATAGGCTCAGAGCCAC 1203
Db      112 AspProGlnArgSerProArgThrArgGlnAlaGlyTyrAlaLeuGly----- 127
Qy      1202 ATCTGTGGAGGGGGGTGTCAAGGGTAGAGATGTCTCTCCAGAGAGAGTGGGCC 1143
Db      128 -----GlnGly---SerAlaGlyLeuGlySerArgGlyProArgProHisProAla 143
Qy      1142 GGAGACTGTGAGGTG-----GATCATTCCTG---CAGTCTGTGATCAGAGGCC 1095
Db      144 PheGlnValGln-TyrSerAlaArgAsnProGlyCysProArgThrTyrArgArgArg-- 162
Qy      1094 TCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1035
Db      163 ---SerGlyAlaGlnArgGlyHisProProPro-----GlyAlaGlyGlnArgPr 178
Qy      1034 TAGTCAGCTCATTCACACTGACAGCAGAGCGGAAAGCCCGCGGCGCAGAGCTCTCG 975
Db      178 oSerGlyProThrGlyAlaArgProAlaAlaProGlyAlaProGlyThrProAlaAlaPr 198
Qy      974 AGGAGCTCAGTCAGGAACAGC---CCGCGAGGGCTGTGAGAGGGCGCCACTCAGTAG 918
Db      198 oglyProGlyGlyAlaAlaValAlaProSerGlyAlaThrProHisProGlnArgGlySe 218
Qy      917 TCGCCCGCAGAGGGCTCCAGATGAGAGAGCTTGAAGATTGATGTGCACAAACACAGAG 858
Db      218 rGlyProAlaAspProProAlaAlaAlaArgLeuProProGluArgGlnArgProArgLe 238
Qy      857 CCCAGGCTCCCTTGTAGCGAGCACTGCGGGTGGCTGCCAAACAGTCCAGCTCCGCGC 798
Db      238 uProGlnAspLeuAlaAlaAlaGlnArgCysProAlaGlyProProProThrArgSerG 258
Qy      797 AGCGCTGGGGCGCGCGGAGTCCAGCGCT-----GCTCCAGAGCTGGCCCATGCTCG 744
Db      258 yAlaAlaAlaGlnArgThrHisArgArgProProGlyCysProArgSerAlaArg----- 276
Qy      743 CAGTACTCTGCTGAACCCGAGCCGAGATGCACAGGTCACTTGCTTCAGAGAGGT 684
Db      277 -----AsnProGlyCysProArgThr----- 283
Qy      683 CTGCGCGGGCTGTGACTGTGCT-----GGGATGCGGCTGGGGCG 645
Db      284 -TyrArgArgArgSerGlyAlaGlnArgGlyHisProProProGly-AlaGlyGlnArgP 303
Qy      644 CCTTCCTCGCGCGCTGTGAGCAGCAGCAGTGGCGG---GCCCG---ACTCCGCGCTCG 591
Db      303 roSerGlyProThrArgGlyArgArgProAlaAlaProGlyAlaProGlyThrProAlaAlaP 323
Qy      590 CGCTTGCTGGGGGGGAGCTGTCTCCA----- 562
Db      323 roGlyProGlyGlyAlaAlaValAlaProSerGlyAlaThrProHisProGluArgGlys 343
Qy      562 ----- 562
Db      343 etGlyProAlaAspProProAlaAlaAlaArgLeuProProGluArgGlnGluProArgL 363

```

```

Qy      561 -----CTGACCTGTGAGAAATTTCGAAATCTGCTTGACTCCCGAGCGGACCG 513
Db      363 euProGlnAspLeuAlaAlaGlnArgCysProAlaGlyProProProThrArgSerG 383
Qy      512 CAGCTACCTC----- 502
Db      383 lYAlaAlaAlaGlnArgThrHisArgArgProProGlyCysProArgSerAlaArgAsnP 403
Qy      501 -----TGTCT-----CTTGAAGAGCTGGA-----G 480
Db      403 roGlyCysProArgThrTyrArgArgArgSerGlyAlaGlnArgGlyHisProProProG 423
Qy      479 GTGCCATAGCTATAGCGTTCTGAGACACTGGCCGCGCGCTTCCGCGCCAGGTGGCGC 420
Db      423 lYAlaGlyGlnArgPro-SerGlyProThrGly-GlyArgProAlaAlaProGlyAlaAr 442
Qy      419 AGCAGGCTGTGGCGGGCGGCGAGCGGAGGAGTATCCCGAGAGCGGAGGTTGCTCG 360
Db      442 oglyThrProAlaAlaProGlyProGlyGlyAlaAlaValAlaProSerGlyAlaThrPr 462
Qy      359 CCGCACTGCGCGCGCGCT---CCAGCTCCAGAGAGCTTAGGCGCTGCGGCGCGG 303
Db      462 oHisProGluArgGlySerGlyProAlaAspProProAlaAlaAlaArgLeuProPro-- 481
Qy      302 GCTAAGCTCCGCGCGCGCCAGAGCTCAT-----CCAGCAGAAAG 261
Db      482 -----GluArgGlnGluProArgLeuProGlnAspLeuAlaAlaGlnArg 497
Qy      260 G---CCAGAGAGCTCCAGCTCGGACTCGGCACTTGGC-----CG 225
Db      497 gCysProAlaGlyProProProThrArgSerGlyAlaAlaAlaGlnArgThrHisArgAr 517
Qy      224 CCAGCA-----CCTGAAACATACGTTGAAC 198
Db      517 gProProGlyCysProArgSerAlaArgAsnProGlyCysProArgThrTyrArgArgAr 537
Qy      197 GACAGCATCCGCTAGTATGTCAGGCACTCATCTCTCCAGACAGGGG----- 149
Db      537 gSerGlyAlaGlnArgGly-----HisProProProGlyAlaGlyGlnArgProSe 554
Qy      148 -----CCGGGGTCAGCCCGGATGAGCGCATTCGCCGGG 117
Db      554 rGlyProThrGlyGlyArgProAlaAlaAlaProGlyAlaProGlyThrProAlaAlaPr 574
Qy      116 GGAGGAGGCGGAGACAGCTCAGAACCCGCGC 86
Db      574 yProGlyGlyGlyAlaAlaValAlaProSerGly 584

```

RESULT 2

Q08B3

BHLFI protein - human herpesvirus 4 (strain B95-8)

C:Species: human herpesvirus 4, Epstein-Barr virus

C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004

C:Accession: A03742

R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.

Mol. Biol. Med. 1, 21-45, 1983

A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus.

A:Reference number: A93065; MUID:85035713; PMID:6092825

A:Accession: A03742

A:Molecule type: DNA

A:Residues: 1-660 <BAN>

A:Cross-references: UNIPROT:P03181

R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; He

Nature 310, 207-211, 1984

A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A:Reference number: A03794; MUID:84270667; PMID:6087149

A:Contents: annotation; protein coding region

C:Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-523

C:Superfamily: human herpesvirus 4 BHLFI protein

Alignment Scores: 8.02e-08 Length: 660

QY	1335	GTGTGTACTCTGTGCCCCCTCCCTTCGTGAAATACAGGCCAGGCCGCCGCCCTCCGAGAGCTAA	1276
Db	99	utleValleuLeuaspargluuapgluugluuaspbuProglySerPro--AlaAl	118
QY	1275	GGGGGACACCTGGGGGTA-----	1249
Db	118	aglySerProvalGlyLeuSerlleArgAlaProSerTherValThreSerSerSerglyPr	138
QY	1248	CTGGGAGAGAAATCACTGCTGGAAATGTTTAGGCTTCAGAGCCCAATCTGTGGAGGG	1188
Db	138	oglyProglyProAlaProAlaProglyProglyYArgArpProArglnHis-----	153
QY	1188	GCTGTCAAGGGGGTAGAGATGGTCGTCCTCCAGAGAAAGTGGGCCGGAACTTGGAGGT	1129
Db	154	-SerglnArglnArgProglyProPro-----	162
QY	1128	GGGATCAATCTCTGCCAGTCTTGATCAGAGAGGCTCTGTGCGGCGCGCCCCCTTCTC	1068
Db	163	-----AlaAlaProglyAlaArgProProProProg	172
QY	1068	CTCCATTCAGACACAGCGCGCGCGCCAGCCTCATATGTACGCTTATCCACATGCACAG	1008
Db	172	lnPro-----ProArgProProProProProAlaProProAlaProProAla	184
QY	1008	CAGCGGAACAGCCTCCGGCCACAGCCTTCGACAGG-----CTCGACAGAA	958
Db	184	laProProAlaProProAlaProAlaArgArgProArgglyAlaSpglyProProArgglyAlYT	204
QY	957	CAGCGCCCGCAGGGGCTCAGACAGGGCGCCACTCAGTAGTCGCCCCAGAAAGCGCTCCAG	898
Db	204	hrArgSerValSerPro-----GlyArgArgArgglyLeuuglyProArgArgHisGln-	221
QY	897	ATAGGAGAGCTGTGAACTGTATGTACCAAAACCAAGAGCCAGAGTCCCTTGAGCGAG	838
Db	222	-----HisSerlnGln-----	225
QY	837	CACCTGCGGTGGCTCCCAACACAGT-----CCAGCTCGCGCCAGCGC	793
Db	226	-----ArgTrpProGlnArgArgHisGlyGlyGlyProLeuProGlnProProProProP	244
QY	792	CTGGG-----GCCGCGGGAGATCCAGCGCTTGCTCCAAAGCTGCGCCATG	748
Db	244	roGlyArgSerArgArgProProAlaAlaAlaAlaProProProAlaGlnGlyThrAlaValV	264
QY	747	CTCGAGATCTCTGCTCGAACCCGAGCGCGGATGTACAGATCATTTGGCTTCAGAGA	688
Db	264	al-----ThreIleThrSerThrAlas	271
QY	687	AGGTGTGCGGGCTGTGACTGTGCTGGGGGTGCGGCTGGGGCCCTTCGCGCGGCTCT	628
Db	271	erProTrpLeuaspbuProAlaAlaAlaAlaArgArgLeuaspPro-----AlaAlaAlaT	289
QY	627	GG-----CACCAACACTG--	615
Db	289	rpArgProGluProArgLeuLeuGlnProGlnLeuGlnLeuGlnLeuHisHisArgA	309
QY	614	-----GGCGGCGCCGACCTCGCGGCT	593
Db	309	rgArgArgAlaArgArgProArgProArgGlyArgglyYArgThrArgProAlaArgG	329
QY	592	GCCGCTTGTTGGGGGAGGAGCTGTGCTCCCACTGACCTGAGAAATTTGCAGAAATGC	533
Db	329	lyArg-----GlyAlaProLeuGlnArgGlnPro-----	338
QY	532	TTGACTGCCGACGGGACGACAGCTAACCTCTGTCTCTTTGAAGAGCTGAGTGCCAT	473
Db	339	-----ArgArgArgArg-----	341
QY	472	AGCTATAGCTTCTTGAGAGACATGCGCCGCGCGGCTTCGCGCCAGGTGCGGACAGAGT	413
Db	342	-----ArgAlaGlyGlnGlyAlaLeuArgArgglyYArgglyYPheserSerSerSers	359

QY 412 CGTGGCGGG-----CCAGACGGCGGAGATTGCCCC----- 380
Db 359 er-GlyGlySerAspSerAspSerAspLeuSerProAlaArgSerProSerAlaProAlaPro 378
QY 379 ---GCAGCGCGAGCTTGCTCTTCGCGCAGCACTGCCCGCGCGCTCCAGCTCCAGCAGAGCT 323
Db 379 AlaAlaAlaAlaAlaAlaAlaAlaArgSerSerSerSerSerSerSerSerSerSer 398
QY 322 CT-----AGGC 317
Db 398 erSerSerSerSerSerSerSerSerGlyGlyGlyAlaAspGlyValArgProGlyAlaP 418
QY 316 CGCTGCGGGCCCGGGCGCTAAGCTCCGCGGGCGCCAGAGCGCTCATCCAGCAGAAAGCCA 257
Db 418 rOLEuAlaArgAlaGlyProProProSerProProAla-ProAlaAlaAlaProAlaPro 437
QY 256 GGAGCT-----CCAGCTCGCAGCTCGG 236
Db 438 SerAlaSerSerAlaSerAlaThrSerSerSerAlaAlaAlaSerProAlaPro 457
QY 235 TCAGTTGCCCCGCCCAACCACTCGAACATACGGTGAAGCAGACATCCGCTAGTACCA 176
Db 458 GluProAlaArgProProArgArg-----LysArgArgSerThr 470
QY 175 GGCACTCAT-----CCTCTCCGACGACGGGGCGCGGGGTGAGC 137
Db 471 AsnAsnHisLeuSerLeuMetAlaAspGlyProProThrAspGlyProLeuLeuThr 490
QY 136 CGGATAGAG-----CAATTCCCGGGGGAG-----GGA 110
Db 491 ProLeuGlyGluProThrProThrProThrAspProProAlaAspGlyArgValArgTyrGly 510
QY 109 GCGGAGACAGCTCAGAACCCGCGCAGATTCTGGCGTCCGATGATCTCTGGCGGAGCGCC 50
Db 511 GlyAlaGlyAspSerArgGlyGly-----LeuTrpAspGluAspAspValArgGlnAla 528
QY 49 GCAGCGC 44
Db 529 AlaAla 530

RESULT 4
CGHUTL
collagen alpha 1(III) chain precursor - human
N:Alternate names: procollagen alpha 1(III) chain
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence revision 01-Sep-1995 #ext change 09-Jul-2004
C:Accession: S05272; S04642; PE0011; S01726; S04887; A90399; A94562; I51868; S59511; A904
R:Prockop, D.J.
Submitted to the EMBL Data Library, February 1989
A:Reference number: S05272
A:Accession: S05272
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1240, 'V', 1242-1466 <PRC>
A:Cross-references: UNIPROT:P02461; EMBL:X14420; NID:g30057; PIDN:CA32583.1; PID:g30058
R:Ala-Kokko, L.; Kontusaari, S.; Baldwin, C.T.; Kuitvanlent, H.; Prockop, D.J.
Biochem. J. 260, 509-516, 1989
A:Title: Structure of cDNA clones coding for the entire prepro-alpha1(III) chain of human
erences.
A:Reference number: S04642; MUID:89350838; PMID:2764886
A:Accession: S04642
A:Molecule type: mRNA
A:Residues: 1-1196 <ALA>
A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CA32583.1; PID:g30058
A:Note: the complete sequence is not shown
R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.
Gene 78, 255-265, 1989
A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene ((
A:Reference number: PE0011; MUID:89378752; PMID:2777083
A:Accession: PE0011
A:Molecule type: DNA
A:Residues: 1-176 <BEN>
A:Cross-references: GB:M26939; NID:g180813; PIDN:AAA52040.1; PID:g180814

R.Toman, P.D.; Ricca, G.A.; de Crombrughe, B.
Nucleic Acids Res. 16, 7201, 1988
A>Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pro
A/Reference number: S01726; PMID:88303360; PMID:3405773
A/Accession: S01726
A/Molecule type: mRNA
A/Residues: 1-170 <TOM>
A/Cross-references: EMBL:X07240; NID:g30060; PIND:CAA30229.1; PID:g30061
A/Note: the authors translated the codon CAG for residue 154 as His
R.Janeczko, R.A.; Ramirez, F.
Nucleic Acids Res. 17, 6742, 1989
A>Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen.
A/Reference number: S04887; PMID:8386015; PMID:2780304
A/Accession: S04887
A/Molecule type: mRNA
A/Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634,
A/Cross-references: EMBL:X15333; NID:g29545; PIND:CAA33387.1; PID:g930045
A/Note: the authors' translation of residues 905-932 is inconsistent with the nucleotide
R.Seyer, J.M.; Kang, A.H.
Biochemistry 16, 1158-1164, 1977
A>Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptide
A/Reference number: A90399; PMID:77134724; PMID:557335
A/Accession: A90399
A/Molecule type: protein
A/Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>
A/Experimental source: liver
A/Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact
R.Seyer, J.M.
submitted to the Atlas, December 1977
A/Reference number: A94562
A/Accession: A94562
A/Molecule type: protein
A/Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>
A/Experimental source: liver
A/Note: author submitted corrections to A90399
R.Milewicz, D.M.; Wiltz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.
Am. J. Hum. Genet. 53, 62-70, 1993
A>Title: Parental somatic and germ-line mosaicism for a multilexon deletion with unusual
fingerprint.
A/Reference number: I51868; PMID:93304430; PMID:8317500
A/Accession: I51868
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 186-194 <MIL>
A/Cross-references: GB:S62325; NID:g386425; PIND:AA19337.1; PID:g4261637
R.Chiodo, A.A.; Silience, D.O.; Cole, W.G.; Bateman, J.F.
Biochem. J. 311, 939-943, 1995
A>Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3
A/Reference number: S59511; PMID:96067614; PMID:7487954
A/Accession: S59511
A/Molecule type: mRNA
A/Residues: 302-423 <CHR>
A/Cross-references: GB:S79877; NID:g1195576; PIND:AA55615.1; PID:g1195577
R.Seyer, J.M.; Kang, A.H.
Biochemistry 17, 3404-3411, 1978
A>Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr pe
A/Reference number: A90414; PMID:79000343; PMID:687591
A/Accession: A90414
A/Molecule type: protein
A/Residues: 359-675, 'N', 677-727 <SEY3>
A/Experimental source: liver
R.Lee, B.; Vitale, E.; Supertl-Furga, A.; Steilmann, B.; Ramirez, F.
J. Biol. Chem. 265, 5256-5259, 1991
A>Title: G to T transversion at position +5 of a splice donor site causes skipping of th
A/Reference number: I55349; PMID:9161621; PMID:1672129
A/Accession: I55349
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 537-605 <LBR>
A/Cross-references: GB:M59312; NID:g180815; PIND:AA55041.1; PID:g180816
R.Seyer, J.M.; Mainardi, C.; Kang, A.H.
Biochemistry 19, 1583-1589, 1980
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from ty

A/Reference number: A90438; PMID:80198282; PMID:6246925
A/Accession: A90438
A/Molecule type: protein
A/Residues: 728-895, 'A', 897-964 <SEY4>
A/Experimental source: liver
R.Cole, W.G.; Chiodo, A.A.; Lalande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Chan,
U. Biol. Chem. 265, 17070-17077, 1990
A>Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping and
A/Reference number: A38303; PMID:91009133; PMID:2145268
A/Accession: A38303
A/Molecule type: mRNA
A/Residues: 861-1015 <COL>
A/Cross-references: GB:U05617; GB:M55603; GB:M59227; NID:g180878; PIND:AA559383.1; PID:g
A/Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syndrome
R.Mankoo, B.S.; Dalgleish, R.
Nucleic Acids Res. 16, 2337, 1988
A>Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.
A/Reference number: S02119; PMID:88189827; PMID:3357782
A/Accession: S02119
A/Status: translation not shown
A/Molecule type: mRNA
A/Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>
A/Cross-references: EMBL:X06700; NID:g30053; PIND:CAA29886.1; PID:g30054
R.Seyer, J.M.; Kang, A.H.
Biochemistry 20, 2621-2627, 1981
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from ty
A/Reference number: A90446; PMID:81208139; PMID:7016180
A/Accession: A90446
A/Molecule type: protein
A/Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-3
A/Experimental source: liver
R.Loidl, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Meyer,
Nucleic Acids Res. 12, 9383-9394, 1984
A>Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollagen
A/Reference number: A93551; PMID:85087944; PMID:6096827
A/Accession: A93551
A/Molecule type: mRNA
A/Residues: 1065-1155, 'P', 1157-1466 <LOI>
A/Cross-references: EMBL:X01655; EMBL:X01742; NID:g29584; PIND:CAA5821.1
R.Miskulin, M.; Dalgleish, R.; Kluge-Beckerman, B.; Renard, S.I.; Tolstochev, P.; Brant
Biochemistry 25, 1408-1413, 1986
A>Title: Human type III collagen gene expression is coordinately modulated with the type
A/Reference number: I52393; PMID:86187804; PMID:3754462
A/Accession: I52393
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1161-1200 <MIS>
A/Cross-references: GB:M13146; NID:g180415; PIND:AA52003.1; PID:g180416
R.Emanuel, B.S.; Caminizaro, L.A.; Seyer, J.M.; Myers, J.C.
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985
A>Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm
A/Reference number: I59025; PMID:85216505; PMID:3858826
A/Accession: I79359
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1165-1196 <EMA>
A/Cross-references: GB:M11134; NID:g180417; PIND:AA52004.1; PID:g180418
R.Chu, M.L.; Weil, D.; de Wet, W.; Bernard, M.; Stippola, M.; Ramirez, F.
J. Biol. Chem. 260, 4357-4363, 1985
A>Title: Isolation of cDNA and genomic clones encoding human pro-alpha1 (III) collagen. Pe
A/Reference number: A92516; PMID:85157600; PMID:2579949
A/Accession: A92516
A/Molecule type: DNA
A/Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>
A/Cross-references: GB:M10615; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB:
A/Experimental source: liver
A/Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given for
action
C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit ((C
3-hydroxy)lated. About 15% of the lysines are 5-hydroxy)lated and some are subsequently O-
C/Gene: GDB:COL3A1
A/Cross-references: GDB:118729; OMIM:120180

A:Map position: 2q31-2q31
 A:Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/1; 1275/1; 1337/3; 1418/3
 A>Note: The list of introns is incomplete; defects in this gene can result in Ehlers-Dan
 C:Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b
 er of their length, is formed with desmosine cross-links made from lysine and allysine
 C:Function:
 A:Description: structural component of extracellular fibrous polymer that maintains inte
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd
 F:1-23/Domin: signal sequence #status predicted <SIG>
 F:123-153/Domin: amino-terminal propeptide #status predicted <PRO>
 F:131-91/Domin: von Willebrand factor type C repeat homology <VMC>
 F:154-1221/Product: collagen alpha 1(I) chain #status predicted <MAT>
 F:154-167/Region: amino-terminal nonhelical telopeptide
 F:168-1196/Region: helical
 F:1091-1093/Region: cell attachment (R-G-D) motif
 F:1197-1221/Region: carboxyl-terminal nonhelical telopeptide
 F:1223-1466/Domin: carboxyl-terminal propeptide #status predicted <CPR>
 F:1238-1466/Domin: fibrillar collagen carboxyl-terminal homology <FCC>
 F:124/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
 F:153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted
 F:154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi
 F:161-1212/Modified site: allysine (Lys) #status predicted
 F:161-284-860-977-1106/Modified site: 5-hydroxylysine (Lys) #status experimental
 F:263/Binding site: carboxylate (Lys) (covalent) #status experimental
 F:584-1094/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental
 F:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental
 F:1106/Binding site: carbohydrate (Lys) (covalent) #status predicted

Alignment Scores:

Pred. No.:	1-59e-07	Length:	1466
Score:	253.50	Matches:	157
Percent Similarity:	30.67%	Conservative:	31
Best Local Similarity:	25.61%	Mismatches:	205
Query Match:	7.39%	Indels:	221
DB:	1	Gaps:	33

US-10-030-271-3 (1-1893) x CGHUTL (1-1466)

```

QY 1526 GGTGTGTGTCGAATT-----AGAGGTCTTGAGAGAAAGCCCTTAGCCAAAGATGT 1473
    |||||      ::      ::|||      ::|||      ::|||      ||
Db 31 GtjCyseerHieuleclYglnserYrAlaAspAArgAspValTlrylsProgluProCys 50
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||

QY 1472 GTGCTGGAGAGCAGCGGCGAGCGCCAGGAAATCTCGACAGGGCGAGTGTGACC 1413
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||
Db 51 GlnIleCyseVal-----CysAspserGlyserValleu 61
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||

QY 1412 ATGAGCGAGTGTGTGACCTGTGACAG-----TGACAGACGCCCAAGATCGAGAT 1362
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||
Db 62 CyAspAspIlelleCyAspAspGlnGluLeuAspCyAspAsnProgluIle----- 79
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||

QY 1361 ACAATGTGACGGGGGCTTTGGTGTGTGTGAGCTGTG---CCCTCCCTCTTGAGATA 1305
    |||||      |||||      |||||      |||||      |||||      |||||
Db 80 -----ProheglYgluCySAspAlaValCyAspProgluInProthrala 94
    |||||      |||||      |||||      |||||      |||||      |||||

QY 1304 CAGGCCAGCCCCCG-----CCTCGAGAGCTTAAGGGGCGACACTCGGGGTAGGA 1254
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||
Db 95 ProthAArgProthProAsnGlyGlnGlyProGlnGlyProlyseGlyAspProgluProPro 114
    ::|||      ::|||      ::|||      ::|||      ::|||      ::|||

QY 1253 GGAGTGTGGAGAAACCTCAGCTGGAATGTTTAGGCTCAGAGCC-----1205
    |||||      |||||      |||||      |||||      |||||      |||||
Db 115 GlyIle---ProglYArgAsnGlyAspProglYIle-ProglYGlnProglYserProgl 133
    |||||      |||||      |||||      |||||      |||||      |||||

QY 1204 -----ACATCTGTGTGGAGGGGCTGCAA-----1181
    |||||      |||||      |||||      |||||      |||||      |||||
Db 133 yserProglYProProglYIleCyseGluSerCySProthnGlyProglInAsnTyserPr 153
    |||||      |||||      |||||      |||||      |||||      |||||

QY 1180 -----GGGTAGAGATGTCGTCCTCCAGAGAGAG 1149
    |||||      |||||      |||||      |||||      |||||      |||||
Db 153 oGlnTyAspserTyAspValIySserGlyAlaIaValGlyGlyLeuAlaGlyTyPr 173
    |||||      |||||      |||||      |||||      |||||      |||||

QY 1148 TGGCCCGGAGACTTGAGAGGTGGATCAATCTGCCAGTCTGTGATAGAGAGGCTCTGTC 1089
    |||||      |||||      |||||      |||||      |||||      |||||

```

```

Db 173 oGlyPro-----175
QY 1088 GGGCGCGCCCTCTCTCTCTCTCATGCAGACAGCGCGCGCCGACCTCATATGCA 1029
    |||||      |||||      |||||      |||||      |||||      |||||
Db 176 ---AlaGlyProProglYProProglYProProglYlYhrSerGlyHAspGlySer-- 193
    |||||      |||||      |||||      |||||      |||||      |||||

QY 1028 GCTCATTCACACTCAACAGCAGCGGAAAGCTTCC-----GGCCC 987
    |||||      |||||      |||||      |||||      |||||      |||||
Db 194 -----ProglYserProglYTyGlnGlyProProglYgluProglYGlnAlaGlyPr 211
    |||||      |||||      |||||      |||||      |||||      |||||

QY 986 ACAGCTCTCGAGCGGCTCAGTCAAGAACAGCCCGCAGGGGCTCGACAGCGCGCA 927
    |||||      |||||      |||||      |||||      |||||      |||||
Db 211 oSerGlyProProglYProProglYAlaIleGlyProSerGlyProAlaGlyLysAspGl 231
    |||||      |||||      |||||      |||||      |||||      |||||

QY 926 CTCAGTAGTGTGCCCCAGAGCGCTTCAGATAGAGAGCTTGAGAACTTGATGTACAA 867
    |||||      |||||      |||||      |||||      |||||      |||||
Db 231 yGluSerGlyArgPro---GlyArgProglYgluArgGlyLeu-----244
    |||||      |||||      |||||      |||||      |||||      |||||

QY 866 ACCACAGAGCCAGAGTCTCTTGAAGCAGACACTGGGCTGCGCCAAACAGCTCAGC 807
    |||||      |||||      |||||      |||||      |||||      |||||
Db 245 -----ProglYPro-----ProglYIleLysGlyProAl 254
    |||||      |||||      |||||      |||||      |||||      |||||

QY 806 TGGCGGCGCA-----797
    |||||      |||||      |||||      |||||      |||||      |||||
Db 254 aGlyIleProglYPheProglYMetLysGlyHAspGlyPheAspGlyArgAsnGlyGl 274
    |||||      |||||      |||||      |||||      |||||      |||||

QY 797 -----797
    |||||      |||||      |||||      |||||      |||||      |||||
Db 274 ulysGlyGluThrGlyAlaProglYLeuLysGlygluAsnGlyLeuProglYgluAsnGl 294
    |||||      |||||      |||||      |||||      |||||      |||||

QY 796 -GGCTGTGGCGCGCGCGGATGCCAGCCCTCTCCAAAGCTGCGCCCATGCTCGCAATAC 738
    |||||      |||||      |||||      |||||      |||||      |||||
Db 294 yAlaProglYProMetGly---ProArgGlyAlaPro-Gly-----306
    |||||      |||||      |||||      |||||      |||||      |||||

QY 737 TCTGTCAACCCGAGCGCGATGTCACAGGCTCTTGCTCAGAGAAAGTGTGGCG 678
    |||||      |||||      |||||      |||||      |||||      |||||
Db 307 -----GluArgGlyArgProglYLeuProglYAlaAlaGlyAlaArgGlyAsnAspGlyA 325
    |||||      |||||      |||||      |||||      |||||      |||||

QY 677 GGTGTGACTGTGTGTGGGTGGGTGGGCGCTCTCCGCG---CCGTGTGCACCA 621
    |||||      |||||      |||||      |||||      |||||      |||||
Db 325 la-----ArgIyserAspGlyGlnProglYProProglYProProglYlYnra 341
    |||||      |||||      |||||      |||||      |||||      |||||

QY 620 CC-----ACTGGCGCGCGCCGACTCG---598
    |||||      |||||      |||||      |||||      |||||      |||||
Db 341 laGlyPheProglYserProglYAlaLysGlygluAlaGlyProAlaGlyserProglYs 361
    |||||      |||||      |||||      |||||      |||||      |||||

QY 597 -----CCGTGCGCGCTGTGGGGGAGGCTG--569
    |||||      |||||      |||||      |||||      |||||      |||||
Db 361 eAsnGlyAlaProglYlYnArgGlygluProglYProGlnGlyHAspAlaGlyAlaGln 381
    |||||      |||||      |||||      |||||      |||||      |||||

QY 568 --TCTCCCACTGACCTGTGTAATTTGCAGAACTGTGACTGCGACGCGAGCGCA 511
    |||||      |||||      |||||      |||||      |||||      |||||
Db 381 lyProProglYProProglYIleAsn-----GlyserProglYglYsGlyg 397
    |||||      |||||      |||||      |||||      |||||      |||||

QY 510 GCTAC---CCTGTGTCCTTTGAAGAGCTGAGAGTGCATAGCTTAGCTTTCGAGA 454
    |||||      |||||      |||||      |||||      |||||      |||||
Db 397 lUmetGlyPro-----400
    |||||      |||||      |||||      |||||      |||||      |||||

QY 453 CACTGCGCGCGCGCTTGCGCGCAGAGTGCAGAGAGTGTGAGC-----GGGCGAG 400
    |||||      |||||      |||||      |||||      |||||      |||||
Db 401 -----AlaGlyIleProglYAlaProglYLeuMetGlyAlaArgGlyProProglYProA 419
    |||||      |||||      |||||      |||||      |||||      |||||

QY 399 CAGCGCAGAGAGTTCGCCAGAGCGCGAGGTTGCTCT---CGCGCACTG 352
    |||||      |||||      |||||      |||||      |||||      |||||
Db 419 laGlyAlaAsnGlyAlaProglYLeuArgGlyGlyAlaGlygluProglYlysAsnGlyA 439
    |||||      |||||      |||||      |||||      |||||      |||||

QY 351 CCGCGGCGGCTCAGCTCAGCAGAGAGCTTAGCGCGCTGCGGG---CCGCGGCTAAGCC 295
    |||||      |||||      |||||      |||||      |||||      |||||
Db 439 laLysGlygluProglYProArgGlygluArgGlygluAlaGlyIleProglYValProG 459
    |||||      |||||      |||||      |||||      |||||      |||||

QY 294 TCCGCGCGCGCAGAGAGCTATCCAGCAGAAAGCGCAGAGCTGCAGCTGCAGCTCGGT 235
    |||||      |||||      |||||      |||||      |||||      |||||
Db 459 lyAlaLysGlygluAspGlyLysAspGlyser---ProglYgluProglY-----474
    |||||      |||||      |||||      |||||      |||||      |||||

```

QY 234 CAGTTGCCGCCCAACCTCGAACATACGGTGAAGGACAGATCCCTGATGATCCAG 175
 Db 475 -----AlaAenGlyLeuProGlyAlaAlaGlyLu-----ArgGlyAlaProG 489
 QY 174 GCACTCATCTCTCTCC-----AGCACGGGGCGGGGCTGCACC 136
 Db 489 LpNeatrgGlyProAlaGlyProAenGlyLeProGlyGluLpGlyProAlaGlyLuA 509
 QY 135 GGATAGGCGCATTCCTCCGGGGAGGAGGACACAGTCAACCGCGCATTCCTCG 76
 Db 509 rGgGlyAlaProGlyProAlaGlyProAlaGlyAlaAlaGlyLuProGly-ArgApgGly 528
 QY 75 CGTCCGATGACTCTCGGCGGACGCGCGACCCA 43
 Db 529 ValProGlyGlyProGlyMetArgGlyMetPro 539
 RESULT 5
 A45344
 Immediate-early protein - suid herpesvirus 1 (strain Kaplan)
 C:Species: suid herpesvirus 1
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C:Accession: A45344
 R:Vicer, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.
 Virology 179, 365-377, 1990
 A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented or
 A:Reference number: A45344; PMID:91021039; PMID:2171211
 A:Accession: A45344
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1446 <VLC>
 A:Cross-references: UNIPROT:P33479; GB:M34651; NID:G334070; PID:AAA47470.1; PID:G334071
 C:Superfamily: herpesvirus immediate-early protein IE175
 C:Keywords: DNA binding; early protein; transcription regulation
 Alignment Scores:
 Pred. No.: 5e-07 Length: 1446
 Score: 245.00 Matches: 155
 Percent Similarity: 33.82% Conservative: 31
 Best Local Similarity: 28.18% Mismatches: 174
 Query Match: 7.14% Indels: 190
 DB: 1 Gaps: 28
 US-10-030-271-3 (1-1883) x A45344 (1-1446)
 QY 1290 GCTCCCGAGGCTTAAGGCGGACACCTGGGGTGAAGAGTCTGGAAAGAACTCAGC 1231
 Db 31 AlaSerGlyProApgGlyGlySerGlnGlySerArgArg-----Arg-GlySerSe 47
 QY 1230 TGGAAATGTTTA-----GGCCTC-----AG 1210
 Db 47 rGlyGlyLpApgLeuNeuNeuGlyProGlyGlyLpNeuPheSerApgAlaAlaGlyLuAgl 67
 QY 1209 AGCCCATCTGTGGAGGGGCTGTCAGAGGGTAGAGTGTCTCTCCACAG----- 1155
 Db 67 uAlaAlaValLeuAlaAlaAlaAlaGlyAlaThrArgProArgProProSerAlaG 87
 QY 1154 -----GAAAGTGGCCCGGAGACTTGGAGTGGGATCAACCTCCAGT 1111
 Db 87 nGlnGlnGlnLpApgArgArgGlySerGlyLu-----LleValVa 101
 QY 1110 CTTGGATCAGAGGCTCTGTGCGGGCGCGCCCTTCTCTTCATCAGACAGAGCG 1051
 Db 101 lLeuApgApgLpLpApgLpLpGlnGlnApgLpLpSer-ProAlaAlaGlySerProG 121
 QY 1050 GC-----GCCGGCAGGCTCATAGTACG-----CTTCATCCAGACT 1015
 Db 121 lYAAGAlaLeuHleGlnGlySerGlnHleGlyHleValLeuGlyProArgSerArg 141
 QY 1014 GACACAGAGGAGAACAGCTCCCGGCGGACAGCTTTCGACGGGCTCAGTCAGAAAC 955
 Db 141 lagGlySerGlyProArgProProThrProAlaAlaLeuAlaAlaGlyLuAglAla- 160

QY 954 GCCCGAGGAGCCTGCAGCAGGGCGCCTACAGTACTGCCCCAGAGGCTCCAGATA 895
 Db 161 --ProGlyGlyProGlyArgSerSerProSerAlaAlaSerPro----- 174
 QY 894 GAGAGCTTGAGAACTGATGTCACAAACACAGAGCCAGGTCCTTACGCGACAG 835
 Db 175 -----AlaSerSerSerGlySerProGlyProSerAlaAlaProA 188
 QY 834 TGGCGTGG-----CTTGCCCAACACAGTCCAGCTCCGCGGCGGCGGCGGGA 778
 Db 188 rGArgTtPserProAlaArgGlyApgProValGlyLpGlnProGlyProAlaAla--ArgPr 207
 QY 777 TGGCAGCGCCTGCTCCAGGCTGCGCCATGCTCGACACTTCTTGAAACCGGAGCC 718
 Db 207 oArgThrProAlaProProAlaGlnProAlaAlaAlaAlaAlaProAlaArgA 227
 QY 717 GATGTACAGACTACTTGGCTTCAGAGAGAGGTCTGGCGGCTTACCTGCTGCGG 658
 Db 227 yProAlaSerProAlaSerProAla-----AlaGlyProValSerAlaProG 243
 QY 657 TGGCGTGGGGCCCTCTC-----CGCGCGCTGTGGACACCACTGGGCGG 610
 Db 243 yGlyGlyGlyAlaProSerAlaGlyApgApgArgGlyArg-HisHisHleGlnHisArg 263
 QY 609 GCCCC-----GACTCCGCGCTGCGCTGCTGTTGGG-----GGAGCC 571
 Db 263 lPProLeuLeuApgLpProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 283
 QY 570 TGTCTCCC-----ACTGACCTGCTGAGAAATTGGCAACTGCTTACTGCC----- 524
 Db 283 rGserProValaSerSerApgApgApgApgApgApgApgApgApgApgApgApgApg 303
 QY 523 -----GAGCGGCGGCGGACACTACCTCTGCTCTTGA 490
 Db 303 luthrValAlaAlaArgGlyProGlnLpApgLpApgLpLeu----- 316
 QY 489 AGAGCTGAGAGTCCATAGCTATAGCTTGTGAGACACTGCGCGCGCTTCGCGC 430
 Db 317 -----GlyLeuAlaGlyApgGlyGlyAla 325
 QY 429 CA-----GTCGCGGACGA-----GTTGTG 409
 Db 325 rProGlnArgLpProArgArgArgArgAlaGlyGlnGlyAlaLeuArgArgGlyArg 345
 QY 408 GC-----GGGCGACGCGGCGAGAG 388
 Db 345 lPheSerSerSerSerSerSerGlySerApgSerApgLpSerProAlaArgSerPro 365
 QY 387 TTGCCCCA-----GAGCGGAGGTTGCTCTGCGGCACTGCGCGGCGCTTCAGCTCAG 331
 Db 365 eAlaProAlaGlnProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 384
 QY 330 CAGGAGCTT----- 321
 Db 384 rSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 404
 QY 320 -----AGGCGCTGCGGCGCGGCTTAAGCTTCGCGGCGGCGGACAGAGCTC 274
 Db 404 yValArgProGlyAlaAlaProLeuAlaAlaAlaAlaGlyProProProProSerProAla- 424
 QY 273 ATCCAGCAGAAAGGCGCAGAGCT-----CAGCTC 244
 Db 424 lAlaAlaAlaProGlnProSerAlaSerSerAlaSerSerAlaAlaAlaAlaAlaAla 444
 QY 243 GCACTGGGTAGTTGCCCGCCACACACTTGAAACATACGTTGAAGCAGCATCCGTA 184
 Db 444 rAlaProGlnProAlaArgProProArgArg-----LysArg 457
 QY 183 GTAGTCAGGACACTCAT-----CCTCTCCAGCAGGCGGCGG 145
 Db 457 rGserThrTrApgApgApgApgApgApgApgApgApgApgApgApgApgApgApgApg 477
 QY 144 GGTGACCCCGATAGCG-----CAATTCGCGGCGGAG----- 113

[illegible]

Db	976	--ProProProProProProProPro-----	ProProProProLaPro	988
QY	810	CAGCTGCGCGCAGCGCCTTGAGCGCCGCGGAGT-----		776
Db	988	roLaProLaProLaProLaProSerSerSerGlyGlyTySerSerGlyGlySerSerGlyGlySerA	1008	
QY	776	-----		776
Db	1008	LaLaGlyGlyGlyGlyTySerSerSerGlyGlyTySerSerGlyGlySerA	LaLaProPro	1028
QY	775	-----CCACGCCCTGCTCCCAAGCGTGGCCCATGCTGCAGATCTCTCTGCAACCCGGAG	721	
Db	1028	roProProProProProProProLa--ProLaProLa-----	ProLa	1042
QY	720	CCGATGTCACAGGTCACTTTGACCTTCAGAGGAGTCTGGCGGCTCTGACTGCTCTG	661	
Db	1043	ProSerSerSerGlyTyTySerGly-----GlySerSerSerGly-----	SerA	1058
QY	660	GGTGTCGGCTGGGGCCCTCTCCGCCCGCTTGCGCACACCACTGGAGCCGCCCACT	601	
Db	1059	GlyGlyGlyGlyGlyGlySerSerSerGlyGlyTySerSerGlyGlySerThraLaProProPro	1078	
QY	600	CCGCGCGCTGCGCGCTTGTTGGGGGGGAGCCGTCTCCACTGACCTCGTAGAATTTC	541	
Db	1079	ProProProPro-----		1082
QY	540	AGAACTGCTTGACTGCCGACGCGCAGCGACGTAACCTCTGTCTCTTTGAGACTGA	481	
Db	1083	-----ProProProProProLaProLaProLaProLaProLaProSerSerSerGly	1098	
QY	480	GGTGCCATAGCTATAGCGTTCTGAGACACTGCCGCGCGCTTGCGCCGACGTGCG	421	
Db	1099	Gly--TySer-----	GlyGlySerSerGlyGlySerA	1109
QY	420	CAGACGCTGTCGGGGCGCAGCGACGCGAGT-----		386
Db	1109	LaLaGlyGlyGlyGlyGlySerSerSerGlyGlyTySerSerGlyGlySerA	LaLaProPro	1129
QY	385	-----	GCCCGACGACCCGAGTT	367
Db	1129	roProProProProProProProLaProLaProLaProLaProLaProSerSerSerGly	1149	
QY	366	GCTTCGCGCAGCTGCGCGCGCGCTCAGCTCAGCAGAGACTTGAGCGCGCTGCGG	308	
Db	1149	yrSerGlyGlySerSerSerGlyGlySerA	LaLaGlyGlyGlyGlyGlyGlyGlySerSerSerGlyG	1168
QY	307	-----CCCGGCTMAAGCTCCGGGGGGCGCAGAGCCTCATCCAGAGAAAGCCAGAG	253	
Db	1168	lyTySerSerGlyGlySerA	LaLaProProProProProProProProProProLa	1188
QY	252	CTTCAGCTCGCACTCGGTCACTTCGCCGCCACCACTCGAATCGGTGAGCGACAG	193	
Db	1188	LaProLa-----	ProLaProLaProSerSerSerGlyGlyTySerSerGlyG	1203
QY	192	CATCC-----		188
Db	1203	lySerSerGlyGlySerA	LaLaGlyGlyGlyGlyGlySerSerSerGlyGlyTySerSerGlyG	1223
QY	187	--CGTAGTAGCCAGGACACTATCTCTCTCCACACAGGGGCGGGGTCCACCGGATAG	130	
Db	1223	lySerA	LaLaProProProProProProProProProProLaProLaProLaProLa	1243
QY	129	CGCAATTCGCGGGGAGGAGCGAACAAGCTCAGAACCCGCGC	86	
Db	1243	LaProLaProSerSerSerGlyGlyTySerSerGlyGlySerSerGly	1257	

RESULT 7

CGCHS

collagen alpha 1(I) chain - chicken (tentative sequence) (fragments)

C:Species: Gallus gallus (chicken)

C:Date: 12-Aug-1981 #sequence_revision 06-Jul-1982 #text_change 31-Mar-2000

C:Accession: A90458; A90181; A02857

C.Species: Strongylocentrotus purpuratus (purple urchin)
 C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C.Accession: S23809
 J.Expositor: J.Y.; d'Alessio, M.; Solursh, M.; Ramirez, F.
 R.Biol. Chem. 267, 15559-15562, 1992
 A.Title: Sea urchin collagen evolutionarily homologous to vertebrate pro-alpha-2(I) coll
 A.Reference number: S23809; PMID:92348411; PMID:1639795
 A.Accession: S23809
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-1414 <EXP>
 A.Cross-references: UNIPROT:Q26634; EMBL:M92040; NID:G16143; PID:AAA30035.1; PID:G1614
 C.Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology
 C.Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
 F.1207-1414/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:

Pred. No.:	1,93e-06	Length:	1414
Score:	235.00	Matches:	175
Percent Similarity:	30.38%	Conservative:	34
Best Local Similarity:	25.44%	Mismatches:	199
Query Match:	6.85%	Indels:	281
DB:	1	Gaps:	37

US-10-030-271-3 (1-1883) x S23809 (1-1414)

```

QY 1678 AGCTCTGACAGCAGCGG-----GGAGGTGAGAGAGTCAAGTCTGTAA 1634
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 481 SerAlaGlySerIleSerIleSerAlaGlyGlnSerGlyIleProGlyIleuProMetPro 500
QY 1633 ATTGAAGCTGG-----CAGGCG-----CCTGGCTGG 1607
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 501 GtGlnArgGlyLeuThrGlyProProGlyIleGlnGlyArgAspGlyLysProGlyPro 520
QY 1606 CTGGAGATGTCGGCAA---GGTGACAGCGCCCGCATGACCCAGCTCCAGTCCCA 1550
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 521 AlaGlyAlaProGlyIleuProGlyIleuSerGlyProAlaGlyAlaSerGlyIleuGly 540
QY 1549 CTGATGTGCTGACAGCCAGGTGTGTGTGTCMAATTAGAGTCTCTGAGAGAACCC 1490
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 541 Leu-----ProGlyLeuValGlyLeu-----ProGlyPro 550
QY 1489 TTAGGCCAAAGAGTGTGTGCTGGGAGACAGCCAGGCGCAGGCGCAGGGA----- 1438
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 551 GlnGlyGlnArgGly-----GlnArgGlyGlnAspGlyGly-GlnGlySerProG 567
QY 1437 -----AGATCTCAGCAGCGGCGAGTGTGATGACATGACGAGTGTGCACTGT 1391
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 567 yAlaProGlyLeuThrGlyGlnProGlyIleuSerGlyIleuProGlyAlaAlaGlyProPr 587
QY 1390 GACAGTCAGACAGCCCAAGATCAGATACATATGTGACAGGCGGCGCTTGTGTGACTGTG 1331
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 587 ogIyProGlnGlySerAlaGlyIleuArgGlyIleu-----GlnGlyProGlnGlyIleuAlaG 606
QY 1330 TTAGCTGCCCCCTCTCTGAGATACAGGCGCCAGCGCCCGCTCCGAGAGGCTAAGGG-- 1273
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 606 ySerMetGlyProPro-----GlyProProGlyAlaSerGlyAlaSerGlyAla 622
QY 1273 ----- 1273
DB 622 agIlnGlyAspAsnGlyProProGlyIleuSerGlyProGlnGlyProProGlyAlaArgG 642
QY 1272 -----GGCAGCCTGG-----GGTAGAG 1253
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 642 yGlnArgGlyAlaProGlyIleuArgGlyProGlnGlyLeuThrGlyAlaGlnGlyArgR 662
QY 1252 GAGCTCGGAGAGAACTCAGCTGCGAAAGTTTAGGCTCAGAGCCACATCTCTGGG 1193
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 662 gGlyIleGly-GlyArgAlaGlyAsnGlyAla-ThrGlyProProGlyIleuIleGly 681
QY 1192 AGG-----GGCTGTCAGAGGGGTAGAGATGTCGTCCTCCAGAGA 1154
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 682 GluMetGlyProProGlyAsnValGlyLeuGlnGlyPro-----ProGly 696
  
```

```

QY 1153 GAAGTGGCCCGAGACTTGGAGGTGGATCATCTCCAGTCTGTGATCAGAGGCTT 1094
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 697 GluLeuGlyPro----- 700
QY 1099 CTGTGGGGGGGGCCGCCCTTCTCTCATCAGCAACAGGCGGCGCGAGCTCAT 1034
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 701 ---SerGly-----ProProGlyAlaArgGlyProProGly---ProSer 713
QY 1033 AGTCAGCTCATCATCAGACTGACAGCAGAGCGAAGC----- 998
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 714 GlySerProGlyProAspGlyProAlaGlyAlaGlnGlyAspArgGlyProValGlyPro 733
QY 997 -----CTTCCGCGCCACAGCTTCCGAGGCGCTCAGTCAAGAACAGCCCGCAGGG 944
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 734 MetGlyProSerGlyProSerGlyMetProGlyIleuArgGlyAspAsnGlyGlnProGly 753
QY 943 CTGACAGCAGCGGCGCAGCTGAGTGTGCGCCAGAAAGCGCTCAGATAGAGAGCTCTG 884
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 754 Pro-----GlyGlyAlaProGlyIleuArgGlyAspProGlyAspLeu 767
QY 883 AGAATTGATGTCACAAACACAG--AGCCAGGTCTCTTGAGCGCAGCACTGCGGTGG 827
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 768 -----GlyProGlnGlySerProGly----- 774
QY 826 CCTGCCAAACAGTTCAGCTGCGCGCCAGCGCTGCGGCGCGCGAGTCCAGCCCT 767
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 775 -----SerProGlyPheAlaGlyProProGlyArgSerGlyAsnProGlyPro 790
QY 766 GCTCCAGGCTGCCCATGCTGCAGTACTGCTGCAACCCGAGCGCGGATGTACAGG 707
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 791 GlnGlyLeu-----Gly-ProThrG 798
QY 706 TCACTTTGCTTCAGAGAGGTGTCGGCGGCTGACTGCTGTGCGGTGCGGTGGG 647
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 798 Y-----AlaArgGlyGlnThrGlyGly----- 805
QY 646 CCCCTCTCCGCGCGTGTGACCAACACAGCGCGCGCGCGCTCCGCGCTG----- 592
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 806 -----ProGlyProSerGlyProThrGlyAspProGlyProGlnGlyProLeuGlyAl 823
QY 591 -----CCGTTGGTGGGGGAGCGCG----- 569
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 823 aProGlyGlnGlnGlyGlnArgGlyGlnThrGlyProGlnGlyGlnGlyProProG 843
QY 568 -----TCTCCACTGACCTTGC 552
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 843 yProIleGlySerLeuGlyAlaProGlyAlaGlnGlyProProGlyProThrGlyProSe 863
QY 551 TGAGAAATTTGACAACTGCTGACTGCGCAGCGCAGCGAGTACCTCTGTCTCTTT 492
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 863 rGlyAsnAla-----GlySerProGlyGlnPro----- 872
QY 491 GAAAGCTGAGAGTGCATAGCTATAGCGTTTGAGAGACATGCGCGCGCGCTTGGGC 432
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 873 -----GlyAlaArgGlyG 877
QY 431 GCCAGTGGCGCA-----GCAGTCTGTGGCGCGCGCAGCA-- 398
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 877 uProGlyGlnSerGlySerProGlyGlnProGlyLeuAlaGlyArgThrGlnGlyProSerG 897
QY 398 ----- 398
DB 897 yGlnArgGlyAspLysGlyAsnAspGlyGlnSerGlyProProGlyProProGlyProAl 917
QY 397 -----CGCGCAGAGTTCGCCACAGCAGCGGAGTTGCTCGCGCAGCATCG 351
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 917 aGlyProAlaGlyGlnSerGlyIleuGlyLeuAlaGlyGlySerGlyProArgGlyPr 937
QY 350 CCGCGCGCTCAGCTCCAGCAGAGAGTCTTAGCGCGCTGCGGCGCGC---GGCTAAGCCT 294
    |||:|||||: |||:|||||: |||:|||||: |||:|||||: |||:|||||:
DB 937 ogIyGlyProAlaGlyProProGlyAlaAlaGlySerArgGlyProAlaGlyLysSerG 957
  
```


QY 234 CAGTTGCCCGCCCA---CCACCTCGAACAATACGCTGAAGCAGACA----- 191
 DB 411 IYATGlyLeuProGlyProProGlyThSeriYAsnProGlyAlaLysGlyThProG 431
 QY 190 -----TCCCGTAGTATCCAGGC-----A 172
 DB 431 IYGLuProGlyLysAsnGlyAlaLysGlyAspProGlyAspProGlyLysGlyLysGlyLys 451
 QY 171 CTCATCTCTCTCCAGACACGGGGCCGGGGTCGACCCGGATAGCCGATTCGCGGGGAGG 112
 DB 451 smGlyThrProGlyAlaIArgGlyPro-----ProGlyGluG 463
 QY 111 GAGGC-----GGAACAAGCTCAAAACCGGCCAGATTCGGGCTCCGATGACTCTGGG 58
 DB 463 IuGlyLysArgGlyAlaAsnGlyLysProGlyLysIuAsnGlyAla-ProGlyThrProGly 482
 QY 57 CGACGGCCGACGCACTGTTTGTGATCTTTCGAGACC 19
 DB 483 GluArgGlySerProGlyPheArgGlyLeuProGlySer 495

RESULT 10

J00405
 hypothetical 119.5K protein (uvrA region) - Micrococcus luteus
 N:Alternate names: ORF 1 protein
 C:Species: Micrococcus luteus; Micrococcus lysodeikticus
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 04-Feb-2000
 C:Accession: J00405
 R:Shiota, S.; Nakayama, H.
 M:J. Gen. Genet. 217, 332-340, 1989
 A:Title: Micrococcus luteus homolog of the Escherichia coli uvrA gene: identification of
 A:Reference number: S04781; MID:89364717; PMID:2549377
 A:Accession: J00405
 A:Molecule type: DNA
 A:Residues: 1-1106 <SH1>
 A:Cross-references: EMBL:X15867
 A:Note: this reading frame extends between two stop codons and does not begin with a sta
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology?

Alignment Scores:
 Pred. No.: 3, 85e-06 Length: 1106
 Score: 230.00 Matches: 176
 Percent Similarity: 32.65% Conservative: 47
 Best Local Similarity: 25.77% Mismatches: 256
 Query Match: 6.70% Indels: 205
 DB: 2 Gaps: 39

US-10-030-271-3 (1-1883) x J00405 (1-1106)

QY 1627 GCTGGGACGGGCTGGCTGGCTGAGATGTG-----GGCAGAGTGAAGCAG 1580
 DB 6 AAGlyGlyGlySerGlyArgSerArgAlaCysAlaThrValTyrGlyArgGlyGlyAsp 25
 QY 1579 GCCCGATGGACCCCGACGTCGAG----- 1556
 DB 26 AAlaProGlyAlaArgAlaIArgLeuSerValGlySerProThrGlyGluThrProValSerCys 45
 QY 1555 ---TGC---CACTGATGTGCTGAAGCCAGCGTGGTGTGTGT-----CAA 1515
 DB 46 ProCysProProLeuLeuSerTyrProArgAlaGlnGluLeuLeuHisIleArgLeuLeuArg 65
 QY 1514 ATTAGAAGTCTCTGAGAAGAA---GCCCTTAGGCCAAAGAAAGTGTGTCTGGGA----- 1464
 DB 66 ArgArgGlyProArgGlyArgProArgLeuArgProArgArgArgAlaLeuGlyArgAla 85
 QY 1463 -----AGCAGGCGCAGGCGAGAGCCC----- 1443
 DB 86 GlyProHisArgArgProGly---ArgProArgAlaGlnProGluLysGlyCysGln 104
 QY 1443 ----- 1443
 DB 105 LeuProAlaGlyArgHisGlyArgValHisGlySerValArgLeuGlyGlnValLeuPro 124

QY 1442 -----AGGAAGATCTCAGCAGGGGCACTGTGACATGAGGAGTGT----- 1401
 DB 125 GlyLeuArgHisAspLeuArgArgGlyProAlaAlaLeuArgValAlaLeuLeuLeu 144
 QY 1401 ----- 1401
 DB 145 ArgProHisValProGlyProGlyGlyGlnAlaGlyArgGlyLeuHisArgGlyProVal 164
 QY 1400 -----GTGCACCTGTGACAGTGCAGACAGCCCAAG----- 1371
 DB 165 ProGlyArgValHisArgProGluValHisGlnProGlnProAlaLeuHisSerGlyAsp 184
 QY 1370 ---ATCAGAGATACATATGTCAGGGGG---GCCTTTGGTGAATGTGTAGCTTGCCCTTC 1317
 DB 185 HisHisArgAspLeuArgLeuHisIleAlaProAlaLeuGlyThrCysArgGlyAlaAlaLeu 204
 QY 1316 CCTTGTGAGATACAGAGCC---CAGCCCCCGCTCCGAGAGGCTAAAGGGGACACACTGGGG 1260
 DB 205 ProAlaValAlaArgArgAlaGlyGluProAlaAspProAlaAlaAsp----- 219
 QY 1259 GTAGAGAGTCTGGGAAGAAATCAGCTGGAAATGTGTTAGGCTCAGAGCCCAATC 1200
 DB 220 -----ArgGlyProAlaArgGly-----AlaAlaArgAlaHisPro 231
 QY 1199 CTGTGGAGAGGCTGTCAAGGGGTAGAGATGTCGTCTCTCCAGAGAA---AGTGGGCC 1143
 DB 233 Leu---ProGly---AlaArgAlaArgGlyProArgProGlnGlyArgValArg-GlyPro 249
 QY 1142 GGAGACTTGAGAGTGGATTCATCTCGACAGCTCGATCAGAGAGGCTGTGCGGGCC 1083
 DB 249 oval-----GlnGlyProValHisIleAlaGlyLeuArgArgGlyArgGlyAs 265
 QY 1082 CCGCCCCCTCTCTCTCATCAGCAAGCGCGCGCGCCAGCTCATATGTCAG----- 1028
 DB 265 ArgArgProAlaLeuGlyProAlaArgAlaGluGlyAlaGlnGlyAlaHisIleArgArg 285
 QY 1027 -----CTCATTCACACTGACACAGCGGCAAGACGCTCCCGGCCACACTCTGCC 975
 DB 285 ArgGlyProProArgHis-----GluGlyGlnGluGlnHisArgGlyValArgArgAlaGlyArgGly 303
 QY 974 AGAG-----CCTCATGACAGAAACAGCCCC---CAGAGGCTGTGAGAGGCGCCACTC 924
 DB 303 yGlyAspArgProGluAlaGlyGlyArgProArgArgGlyArgValArgGlyArgGlyAl 323
 QY 923 AGGTAGTGGCCCCAGAGGCGCTCCAGATAGAGAGCTGAGACCTGATGTCACAAAC 864
 DB 323 acGlyArgArg---GluGlyGlnGluGlnHisArgGlyValArgArgAlaGlyArgGly 342
 QY 863 ACAGAGCCCAAGTCCCTTG----- 845
 DB 342 yGlnProProValProLeuValLeuArgGlyAlaLeuLeuProGlnArgAlaArgAlaAs 362
 QY 844 -----AGCGAGCACTG-----CGGTGGCTGCG 822
 DB 362 ArgArgArgArgArgAlaAlaLeuValLeuLeuGlnGlnProValArgArgValProArg 382
 QY 821 CCAACACGTCCAGCTGCCGCGCAGCGCTGGGGCGCGCGGATGCCACGCGCTGTC 762
 DB 382 GValHisArgHisArgLeuProProAlaGlyLys-----ProGlyPro 396
 QY 761 AAGCTGGGCGCATGCTCGCAGTACTGTGTGCAACCGGAGCGGATGTCAAGACT 702
 DB 396 ArgArgArgGlnArgArgAlaVal-----ProAlaArgGlyArgArgArgAlaVal 413
 QY 701 TTGCTGTAGAGAAAGTGTGCGGGCTGTGACTGTGCTGGGAGTGGGGGCGCT 642
 DB 413 IValAlaArgGlnValHisLeuGlyLeuLeuAlaAlaArgAlaArgArgAlaGlyGln 433
 QY 641 CTCGCGCGCGCTGTGACACACACTGGGCGCGCGCGCGCTGCGCGCTGCTGCTG 582
 DB 433 yAspGlyLeuLeuProGlnHisProValGluGlyProAspArgGlyGlyAlaArgArg 453
 QY 581 G-----GGGGGAGCGCTG---TCTCCCATGACCTGCTGTA 549

Db 453 gProIaLaArgInGlyLeuInGInGlyGlyValArValProGInProValaLaArgProaLa 473
Oy 548 GAATTGCAAGAACTGCTTGACTGCGGACGCGACGCGACGCACTACCTTGTCTCTTGA 489
Db 473 aProLeuInIhIhGlyLeuAaG-----GlyAghIhIhProLeuAaGArLaIaGI 489
Oy 468 GAGC-----TGAGGTGCGCATAGACTTAAGCTTCTG-----GA 456
Db 489 naLaArGlyLaArArGValaGlyAaGArProaLaLaLeuAaGValaInIhIaLaGlyAa 509
Oy 455 GACACTGGCGGCGCGCTTGCGCGCA-----GGTGGCGGACAGCTGCTG 408
Db 509 pProValaProGlyValaProaGlyProProGInProInIhIhProGInaGLaArgGI 529
Oy 407 CGGCGCACACGCGGAGATTGCGGACAGACCGCGAGTTCTCTGCGGCACTGCGCG 348
Db 529 yArGProValaInIhArGlyAaGhIhIhProaLaInIhIaArGlyAaGlyValaLeu 549
Oy 347 CGGCGCTCCAGCTCCAGACGAGACTTAAGCGCGCTGCGGCGCGGCTAAGCTCCGCG 288
Db 549 uGlyValaIaIaLaArGlyProGlyValaIaLaArArGlyProGly-----Al 565
Oy 287 GCGCCAGAGAGCTCATCCGACGAAGAAGCGACAGAGCTCCAGCTCGGACGTTGC 228
Db 565 aGInIhIhArProGlyProaIa-----GlyValaProaLaGlyAaGArProaGlyA 582
Oy 227 CGGCCACACCTCGAACAATACGATGAGAGGACAGATCCCTAGTATGTCAGGCACTCA 168
Db 582 l---ProGInProaGLaIaLaGlyAaGhIhIhProLeuAaGLaGLaGlyPro---AlaIh 600
Oy 167 TCCTCTCTCCAGACAGCGGCGCGGCTGCAACCGGATAGCGCAATTCGCGGAGAGGAG 108
Db 600 sProProGlyInIhIhArArGlyLeuAaGLaGlyAaGArProLeuAaGProaGArGa 620
Oy 107 CGGAACAAGCTCAGAACCCGCGCAGATTGTGCGCTCCGATGACTCTGTGGGCAACGCGCG 48
Db 620 la-----ValhIhArGProaLaProaLaGlyGInPro-----ProProInIhArGaArP 636
Oy 47 AGCCA 43
Db 636 roPro 637

RESULT 11
S59856
collagen alpha 1(III) chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1996 #sequence revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S59856; S62120; S16373
R:Tomam, P.D.; de Crombrughe, B.
Gene 147, 161-168, 1994
A:Title: The mouse type-III procollagen-encoding gene: genomic cloning and complete DNA
A:Reference number: S59856; MUID:95011609; PMID:7926795
A:Accession: S59856
A:Molecule type: DNA
A:Residues: 1-1464 <TOM>
R:Tomam, D.
A:Cross-references: UNIPROT:P08121; EMBL:X52046
submitted to the EMBL Data Library, November 1994
A:Reference number: S62120
A:Accession: S62120
A:Molecule type: DNA
A:Residues: 1-866, 'G', 868-1464 <TOA>
R:Cross-references: EMBL:X52046; NID:9575321; PIND:CAA36279.1; PID:9575322
R:Meersaeranta, M.; Tomam, D.; de Crombrughe, B.; Vuorio, E.
Biochim. Biophys. Acta 1089, 241-243, 1991
A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
A:Reference number: S16176; MUID:91274355; PMID:2054384
A:Accession: S16373
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1442-1464 <MET>
R:Cross-references: EMBL:X57983; NID:950476; PIND:CAA41048.1; PID:950477
A:Cross-references: EMBL:X57983; NID:950476; PIND:CAA41048.1; PID:950477

C:Genetics: 29/1/: 96/3: 112/3: 150/3: 175/3: 193/3: 211/3: 229/3: 247/3: 265/3: 283/3: 296/3: 314/3: 332/3: 350/3: 368/3: 386/3: 404/3: 422/3: 440/3: 458/3: 476/3: 494/3: 512/3: 530/3: 548/3: 566/3: 584/3: 602/3: 620/3: 638/3: 656/3: 674/3: 692/3: 710/3: 728/3: 746/3: 764/3: 782/3: 800/3: 818/3: 836/3: 854/3: 872/3: 890/3: 908/3: 926/3: 944/3: 962/3: 980/3: 998/3: 1016/3: 1034/3: 1052/3: 1070/3: 1088/3: 1106/3: 1124/3: 1142/3: 1160/3: 1178/3: 1196/3: 1214/3: 1232/3: 1250/3: 1268/3: 1286/3: 1304/3: 1322/3: 1340/3: 1358/3: 1376/3: 1394/3: 1412/3: 1430/3: 1448/3: 1466/3: 1484/3: 1502/3: 1520/3: 1538/3: 1556/3: 1574/3: 1592/3: 1610/3: 1628/3: 1646/3: 1664/3: 1682/3: 1700/3: 1718/3: 1736/3: 1754/3: 1772/3: 1790/3: 1808/3: 1826/3: 1844/3: 1862/3: 1880/3: 1898/3: 1916/3: 1934/3: 1952/3: 1970/3: 1988/3: 2006/3: 2024/3: 2042/3: 2060/3: 2078/3: 2096/3: 2114/3: 2132/3: 2150/3: 2168/3: 2186/3: 2204/3: 2222/3: 2240/3: 2258/3: 2276/3: 2294/3: 2312/3: 2330/3: 2348/3: 2366/3: 2384/3: 2402/3: 2420/3: 2438/3: 2456/3: 2474/3: 2492/3: 2510/3: 2528/3: 2546/3: 2564/3: 2582/3: 2600/3: 2618/3: 2636/3: 2654/3: 2672/3: 2690/3: 2708/3: 2726/3: 2744/3: 2762/3: 2780/3: 2798/3: 2816/3: 2834/3: 2852/3: 2870/3: 2888/3: 2906/3: 2924/3: 2942/3: 2960/3: 2978/3: 2996/3: 3014/3: 3032/3: 3050/3: 3068/3: 3086/3: 3104/3: 3122/3: 3140/3: 3158/3: 3176/3: 3194/3: 3212/3: 3230/3: 3248/3: 3266/3: 3284/3: 3302/3: 3320/3: 3338/3: 3356/3: 3374/3: 3392/3: 3410/3: 3428/3: 3446/3: 3464/3: 3482/3: 3500/3: 3518/3: 3536/3: 3554/3: 3572/3: 3590/3: 3608/3: 3626/3: 3644/3: 3662/3: 3680/3: 3698/3: 3716/3: 3734/3: 3752/3: 3770/3: 3788/3: 3806/3: 3824/3: 3842/3: 3860/3: 3878/3: 3896/3: 3914/3: 3932/3: 3950/3: 3968/3: 3986/3: 4004/3: 4022/3: 4040/3: 4058/3: 4076/3: 4094/3: 4112/3: 4130/3: 4148/3: 4166/3: 4184/3: 4202/3: 4220/3: 4238/3: 4256/3: 4274/3: 4292/3: 4310/3: 4328/3: 4346/3: 4364/3: 4382/3: 4400/3: 4418/3: 4436/3: 4454/3: 4472/3: 4490/3: 4508/3: 4526/3: 4544/3: 4562/3: 4580/3: 4598/3: 4616/3: 4634/3: 4652/3: 4670/3: 4688/3: 4706/3: 4724/3: 4742/3: 4760/3: 4778/3: 4796/3: 4814/3: 4832/3: 4850/3: 4868/3: 4886/3: 4904/3: 4922/3: 4940/3: 4958/3: 4976/3: 4994/3: 5012/3: 5030/3: 5048/3: 5066/3: 5084/3: 5102/3: 5120/3: 5138/3: 5156/3: 5174/3: 5192/3: 5210/3: 5228/3: 5246/3: 5264/3: 5282/3: 5300/3: 5318/3: 5336/3: 5354/3: 5372/3: 5390/3: 5408/3: 5426/3: 5444/3: 5462/3: 5480/3: 5498/3: 5516/3: 5534/3: 5552/3: 5570/3: 5588/3: 5606/3: 5624/3: 5642/3: 5660/3: 5678/3: 5696/3: 5714/3: 5732/3: 5750/3: 5768/3: 5786/3: 5804/3: 5822/3: 5840/3: 5858/3: 5876/3: 5894/3: 5912/3: 5930/3: 5948/3: 5966/3: 5984/3: 6002/3: 6020/3: 6038/3: 6056/3: 6074/3: 6092/3: 6110/3: 6128/3: 6146/3: 6164/3: 6182/3: 6200/3: 6218/3: 6236/3: 6254/3: 6272/3: 6290/3: 6308/3: 6326/3: 6344/3: 6362/3: 6380/3: 6398/3: 6416/3: 6434/3: 6452/3: 6470/3: 6488/3: 6506/3: 6524/3: 6542/3: 6560/3: 6578/3: 6596/3: 6614/3: 6632/3: 6650/3: 6668/3: 6686/3: 6704/3: 6722/3: 6740/3: 6758/3: 6776/3: 6794/3: 6812/3: 6830/3: 6848/3: 6866/3: 6884/3: 6902/3: 6920/3: 6938/3: 6956/3: 6974/3: 6992/3: 7010/3: 7028/3: 7046/3: 7064/3: 7082/3: 7100/3: 7118/3: 7136/3: 7154/3: 7172/3: 7190/3: 7208/3: 7226/3: 7244/3: 7262/3: 7280/3: 7298/3: 7316/3: 7334/3: 7352/3: 7370/3: 7388/3: 7406/3: 7424/3: 7442/3: 7460/3: 7478/3: 7496/3: 7514/3: 7532/3: 7550/3: 7568/3: 7586/3: 7604/3: 7622/3: 7640/3: 7658/3: 7676/3: 7694/3: 7712/3: 7730/3: 7748/3: 7766/3: 7784/3: 7802/3: 7820/3: 7838/3: 7856/3: 7874/3: 7892/3: 7910/3: 7928/3: 7946/3: 7964/3: 7982/3: 8000/3: 8018/3: 8036/3: 8054/3: 8072/3: 8090/3: 8108/3: 8126/3: 8144/3: 8162/3: 8180/3: 8198/3: 8216/3: 8234/3: 8252/3: 8270/3: 8288/3: 8306/3: 8324/3: 8342/3: 8360/3: 8378/3: 8396/3: 8414/3: 8432/3: 8450/3: 8468/3: 8486/3: 8504/3: 8522/3: 8540/3: 8558/3: 8576/3: 8594/3: 8612/3: 8630/3: 8648/3: 8666/3: 8684/3: 8702/3: 8720/3: 8738/3: 8756/3: 8774/3: 8792/3: 8810/3: 8828/3: 8846/3: 8864/3: 8882/3: 8900/3: 8918/3: 8936/3: 8954/3: 8972/3: 8990/3: 9008/3: 9026/3: 9044/3: 9062/3: 9080/3: 9098/3: 9116/3: 9134/3: 9152/3: 9170/3: 9188/3: 9206/3: 9224/3: 9242/3: 9260/3: 9278/3: 9296/3: 9314/3: 9332/3: 935

```

QY 881 AACTGATGTCACAAACACAGAGCCAGTCCCTTGAGCGAGCAGCAGTGGGTGCTGC 822
Db 244 -----ProGlyPro-----Pro1 248
QY 821 CCAAACAGCTCGAGTGGCCGCCA----- 797
Db 248 yllelsgslProAlaGlyMetProGlyPheProGlyMetLysGlyHisArgGlyPheAs 268
QY 797 ----- 797
Db 268 pglYArgAnslYgluLysGlygluThrGlyAlaProGlyLeuLysGlygluAsnGlyLeu 288
QY 796 -----GGCGCTGGGGGCGCGGAGATGCAACCGCCCTGTCACAGGTGCG 753
Db 288 uProGlyAspAsnGlyAlaProGlyPheMetGly---ProArgGlyAlaPro-Gly- 305
QY 752 CCATGCTCCAGACTCTGCTCCAAACCCGAGCGGATGTCACAGTTCACATTGCTTCA 693
Db 306 -----GluArgGlyArgProGlyLeuProGlyAlaAlaGlyAlaAla 319
QY 692 GAGGAAGTCTGGCGGCTCTGACTGCTGGGCTGGGCGCGGCTGCTGCGCGG- 634
Db 319 rglYAsnAspGlyAla-----ArgGlySerAspGlyGlnProGlyProProG 335
QY 633 ---CCGCTTGGCACCACC-----ACTGGGCGCG 609
Db 335 lYProProGlyThrAlaGlyPheProGlySerProGlyAlaLysGlygluValGlyProA 355
QY 608 CCCGCACTCG-----CCGCTGCGCGCTTGTT 582
Db 355 lAglySerProGlySerAnslYserProGlyGlnArgGlygluPProGlyProGlnGlyH 375
QY 581 GGGGGGAGCCTG---TCTCCACTGACCTGCTGAGAAATTGACAACTGCTGACTG 526
Db 375 lAlaGlyAlaGlnGlyProProGlyProProGlyAsnAsn-----GlySerP 391
QY 525 CCGACGCGCAGCGGCACTAC---CCTCTGCTCTTGTGAAGAGCTGAGAGTGCATAGCT 469
Db 391 rglYglYlYsglylYMetGlyPro----- 399
QY 468 ATAGCCTTCTGAGACACTGCGCGCGCTTGCGCGCAGAGTGGCGAGCAGTCTG 409
Db 400 -----AlaGlylYleProGlyAlaProGlylYleuGlyAlaArg 413
QY 408 GCGGCGCAGCAGCGCAGAGATT-----GCCCGACAGCCCGCAGTGTCT- 362
Db 413 lYProProGlyProAlaGlyThrAsnGlylYleProGlyThrArgGlyProSerGlygluP 433
QY 361 -----CGCGCACTGCGCGCGGCTCCAGCTCCAGACAGAGTCTAGCGCTGGCGG- 308
Db 433 rglYlYsAsnGlyAlaLysGlygluPProGlyAlaAlaArgGlygluArgGlylYs 453
QY 307 -----CCGGGCTTAGAGCTCCGGGCGCGCAGAGCTTCATCCAGCAGAAAGC 259
Db 453 ePProGlylYleProGlyProLysGlygluAspGlylYsAspGlylYsPProGlyglu- 472
QY 258 CAGGAGCTCCAGTCCGACTCGGTCACTGTCAGTTCGCCCGCCACCACTTCAGACATGCGTGAAG 199
Db 472 rglYlYlYsAsnGlyLeuProGlyAlaAlaGlygluArgGlyProSerGlyPheArgGlyP 492
QY 198 CGACAGCATCCCGTACTAGTCCAGGCACTCATCTCTCCAGCAG-----GGCGCG 145
Db 492 rAlaGlyProAsnGlylYleProGlygluLysGlylYProGlygluArgGlyglYlYProG 512
QY 144 GGTGCAACCGGATAGGCGCATTCGCGGGGAGGAGGCGGAAACAGCTCAGAACCGGCGC 85
Db 512 lY-----ProAlaGlyProArgGlyValAlaGlygluProGly- 524
QY 84 AGATTCTGCGTCCGATGACTCTGGGCGCAGCGCCGACCA 43
Db 525 ArgAspGlyThrProGlyglYProGlylYleArgGlyMetPro 538

```

CG8015
 Collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 24-Apr-1984 #sequence revision 31-Dec-1993 #text change 09-Jul-2004
 C:Accession: A91193; A91229; A91387; A91211; A91201; A43048; A02853
 R:Rauterberg, J.; Timbl, R.; Furthmayr, H.
 Eur. J. Biochem. 27, 231-237, 1972
 A:Title: Structural characterization of N-terminal antigenic determinants in calf and human
 A:Reference number: A91193; PMID:7225534; PMID:415172
 A:Accession: A91193
 A:Molecule type: protein
 A:Residues: 1-19 <RAU>
 A:Cross-references: UNIPROT:P02453
 A:Experimental source: skin
 A:Note: The epsilon carbon of Lys-9, by homology with the rat alpha 1(I) chain, is converted
 R:Fietzek, P.P.; Kuehn, K.
 Eur. J. Biochem. 52, 77-82, 1975
 A:Title: The covalent structure of collagen: amino-acid sequence of the cyanogen-bromide
 A:Reference number: A91229; PMID:76022320; PMID:1164916
 A:Accession: A91229
 A:Molecule type: protein
 A:Residues: 20-145 <FIB>
 A:Experimental source: skin
 A:Note: Lys-103 is hydroxylated and binds glucosylgalactose
 R:Fietzek, P.P.; Mendt, P.; Kell, I.; Kuehn, K.
 FEBS Lett. 26, 74-76, 1972
 A:Title: The covalent structure of collagen: amino acid sequence of alpha1-CB3 from calf
 A:Reference number: A91387; PMID:73049499; PMID:4673951
 A:Accession: A91387
 A:Molecule type: protein
 A:Residues: 146-294 <F12>
 A:Experimental source: skin
 R:Fietzek, P.P.; Rextodt, F.W.; Hopper, K.E.; Kuehn, K.
 Eur. J. Biochem. 38, 396-400, 1973
 A:Title: The covalent structure of collagen. 2. The amino-acid sequence of alpha1-CB7 frc
 A:Reference number: A91211; PMID:74086118; PMID:4359390
 A:Accession: A91211
 A:Molecule type: protein
 A:Residues: 295-562 <F13>
 A:Experimental source: skin
 R:Mendt, P.; Mark, K.V.D.; Rextodt, F.; Kuehn, K.
 Eur. J. Biochem. 30, 163-168, 1972
 A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues.
 A:Reference number: A91201; PMID:73042276; PMID:4343808
 A:Accession: A91201
 A:Molecule type: protein
 A:Residues: 563-675 <MEN>
 A:Experimental source: skin
 R:Fietzek, P.P.; Rextodt, F.W.; Mendt, P.; Stark, M.; Kuehn, K.
 Eur. J. Biochem. 30, 163-168, 1972
 A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6-C
 A:Reference number: A91200; PMID:73042275; PMID:4343807
 A:Accession: A91200
 A:Molecule type: protein
 A:Residues: 676-758 <F14>
 A:Experimental source: skin
 A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in positive
 R:Rauterberg, J.; Fietzek, P.; Rextodt, F.; Becker, U.; Stark, M.; Kuehn, K.
 FEBS Lett. 21, 75-79, 1972
 A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of t
 A:Reference number: A43048
 A:Accession: A43048
 A:Molecule type: protein
 A:Residues: 759-779 <RA2>
 A:Experimental source: skin
 C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydroxy
 C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) are hy
 C:Comment: The order of the eight CNBR peptides in the alpha 1(I) chain of bovine skin c
 9, 149, 268, and 217 residues.
 C:Comment: The complete chain contains 1052 residues.
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglyutamic acid; trimet
 F:1/Modified site: pyroglutamic acid (Gln) #status experimental

Alignment Scores:

Pred. No.: 5.17e-06 Length: 779
 Score: 228.00 Matches: 128
 Percent Similarity: 33.49% Conservative: 16
 Best Local Similarity: 29.77% Mismatches: 181
 Query Match: 6.65% Indels: 106
 DB: 1 Gaps: 21

US-10-030-271-3 (1-1883) x CGBOLS (1-779)

QY 1308 GATACAGGCCCAAGCC-----CCCGCTCCGAGGCTTAAGGGGCAACACT 1264
 DB 117 AspaIagIyProIagIyProIyagIyGIuProIySerProIyGIuAsnGIyAlaPro 136
 QY 1263 GGG-----GGTAGAGAGAGTGTGGGAAGAACTCAGCTGGAATGGTTAGGCTCAG 1210
 DB 137 GIyGIuMetGIyProIagIyLeu-ProGIyPheProGIyProIyagIyAla-AlaGIyG 156
 QY 1209 AGCCCACTCTCTGGG---AGGGCTGTCAAGGGGTAGAGATGGTCTCTCCAGAG 1153
 DB 156 IuProGIyLyAlaGIyGIuIuArgIyValProGIyPro-----ProGIyA 171
 QY 1152 AAGGTGGCCCGAGACTTGGAGGTGGATCAATCTGCCACTCTGGATCAGAGGCTTC 1093
 DB 171 IaValGIyProIaGIyLyAsnAspGIyGIuAlaGIyAlaGIy-----ProP 187
 QY 1092 TGTGGGGGCGCCCGCTCTCTCCATCAGCAACAGGGGCGCGGAGGCTCTATA 1033
 DB 187 roGIyProIaGIyProIaGIyGIuIuArgIyGIuGIyProIaGIySerProGIyP 207
 QY 1032 GTACG-----CTCATCCACACTGACCAAGAGGCGGAGCAAG---CTTCCGCGCCACAGC 982
 DB 207 heGIuGIyLeuProGIyProIaGIyProProGIyGIuAlaGIyLySerProGIyGIuGIy 227
 QY 981 CTCTCGAGGGCTTCAGTCAAGAAACGCGCCGAGGGCTTCAGAGAGGGCGC---CACT 925
 DB 227 IyValIProGIyAspLeuGIyAlaProGIyProSerGIyAlaArgGIyGIuIuArgIyPheP 247
 QY 924 CAGGTAGTGGCCCGCAAGAGGCGTCCAGATAGAGAGCTCTAGAACTTAGATGTCAACAAC 865
 DB 247 roGIyGIuIuArgIyValGIuGIyPro----- 255
 QY 864 CACAGAGCCCAAGTCCCTTGGAGCCGACAGCATCGGTGGCTGCCCAACAGTCAAGTGC 805
 DB 256 -----ProGIyProIaGIyProArg-----GIyAlaAsnGIyAlaProGIyA 270
 QY 804 CCGCGCAGCGCTGGGGCGCCGAGATGCCAG---CCCTGCTCAAGGCTGGCCCAT 749
 DB 270 snAspGIyAlaIyAspIaGIyAlaProGIyAlaProGIySerGIyAlaIyAlaProG 290
 QY 748 GCTGCGAGTACTCTGCTGCAACCGGAGCCGAGTGTCAAGTCACTTGTGCTTCAGAGG 689
 DB 290 IyLeuGIyGIyMetProGIyGIuIuArgIyAlaIaIy-----LeuProIyProL 307
 QY 688 AAGGTCTGGCGGCTGTGACTGTGCTGGGGTGGGCGGCTCTCCG----- 636
 DB 307 yseGIyAspAspAspIyAspIaGIyProIyagIyAlaAspGIyAlaProGIyLyAspGIyV 327
 QY 635 --CGCGCTTGGCAACCACTGGGCGCGCCGAGCTCCGCGCTGGCTTGGTTGGGG 578
 DB 327 alArGIyLeuThrGIyProIleGIy-ProProGIyProIaGIyAlaProGIyAspLyAs 346
 QY 577 GGGAGCGCTGTCCACATGAGCCCTGTGAGAAATTGCAAGACTGCTTACGCGGAGGCG 518
 DB 347 GIyGIuAlaGIyProSerGIyPro----- 354
 QY 517 GACGGCAGCTACCTCTGTCTTTTGAAGAGCTGAGGTGCATAGCTATAGGTTCTG 458
 DB 354 ----- 354
 QY 457 GAGACACTGGCGCGCGCTTGGCGCGCAGGTGCGGAGAGGTCTGTGGCGGCGCAGCA 398
 DB 457 ----- 398

DB 355 -----AlaGIyThrArgGIyAlaProGIyAspArgGIyGIuProGIyProProGIy 371
 QY 397 CGCGCAGAGAGTGGCCCGCAGAGCGCAGAGTGTCTCTGCCGCACTGCCGCGGCGCTCA 338
 DB 372 ProIaGIyPheAlaGIyProProGIyAlaAspGIyGIuProGIyAlaIyAspGIyGIuPro 391
 QY 337 GCTTCAGAGAGAGTCTTAAAGCGCGCGCGCGCGCTTAAAGCTCCGCGGCGCAGAG 278
 DB 392 GIyAspAlaGIyAlaIyAspIaGIyPro----- 402
 QY 277 CCTATCCAGCAAGAAAGCCAGAGCTCCAGCTCCAGCTCGTCACTTGTGCTCCGCCACCA 218
 DB 403 ProGIyProIaGIyProIaGIyProProGIyProIleGIyAsnValGIyAlaProGIy 422
 QY 217 CCTGCAACATACGCTTGAAGCAGACATCCCGCTGTAGTGTCCAGGCACTCATCTCTCC 158
 DB 423 ProLyS-----GIyAlaArgGIySerAlaGIy-----ProProGIy 434
 QY 157 AGCAGGGG---CCGGGTGACAC---CGATACGCCCATTCGCGGGGAGGAGGC--- 107
 DB 435 AlArThrGIyPheProGIyAlaIaGIyArgValGIyPro---ProGIyProSerGIyAsn 453
 QY 106 ---GGAACAAGCTCAAGAACCCGCGC 86
 DB 454 AlaGIyProProGIyProProGIy 461
 RESULT 13
 S21626
 collagen alpha 1(I) chain precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence revision 25-Apr-1997 #text change 09-Jul-2004
 C:Accession: S57243, J16374, A23982, I49559, I49579, I48300, S21626
 R:IL, S.W.; Killian, J.; Prockop, D.J.
 A:Title: The complete cDNA coding sequence for the mouse pro-alpha-1(I) chain of type I
 A:Reference number: S57243
 A:Accession: S57243
 A:Molecule type: mRNA
 A:Residues: 1-1453 <LIS>
 A:Cross-references: UNIPROT:P11087; EMBL:U08020; NID:G470673; PIDN:AAA88912.1; PID:G47067
 R:Metzger, M.; Tomar, D.; de Crombrughe, B.; Vuorio, E.
 A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
 A:Reference number: S16176; MUID:9127435; PMID:2054384
 A:Accession: S16176
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1442-1453 <MET>
 A:Cross-references: EMBL:X57981; NID:G50484; PIDN:CAA1046.1; PID:G50485
 R:French, B.T.; Lee, W.H.; Maul, G.G.
 A:Title: Nucleotide sequence of a cDNA clone for mouse proalpha1(I) collagen protein.
 A:Reference number: A23982; MUID:86137403; PMID:3841523
 A:Accession: A23982
 A:Molecule type: mRNA
 A:Residues: 518-1128 <PRE>
 A:Cross-references: GB:M17491; NID:G192261; PIDN:AAA7333.1; PID:G192262
 R:Monson, J.M.; Friedman, J.; McCarthy, B.J.
 A:Title: DNA sequence analysis of a mouse pro-alpha-1(I) procollagen gene: Evidence for
 A:Reference number: I49559; MUID:83141374; PMID:6298597
 A:Accession: I49559
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 735-1130 <RES>
 A:Cross-references: GB:M17491; NID:G192263; PIDN:AAA7334.1; PID:G192264
 R:Haebler, K.; Kuehn, M.; Delius, H.; Jaenisch, R.
 A:Title: Acad. Sci. U.S.A. 81, 1504-1508, 1984
 A:Reference number: I49557; MUID:84170331; PMID:6324198
 A:Accession: I49557
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-25 <RR2>
 A:Cross-References: GB:K01688; NID:g192246; PIDN:AAA37330.1; PID:g553881
 R:Pencon, S.P.; Lamande, S.R.; Hannagan, M.; Stacey, A.; Jaenisch, R.; Bateman, J.F.
 Biochem. Biophys. Acta 1216, 469-474, 1993
 A:Title: Genomic sequence of mouse COL1A1 encoding the collagen propeptides.
 A:Reference number: S33789; MID:94092741; PMID:8268229
 A:Accession: S33789
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-185, 1031-1201, 'G', 1203-1218, 'E', 1220-1221, 'T', 1223-1
 R:Rhodes, K.; Rippe, R.A.; Umezawa, A.; Nehls, M.; Brenner, D.A.; Breindl, M.
 Mol. Cell. Biol. 14, 5950-5960, 1994
 A:Title: DNA methylation represses the murine alpha 1(I) collagen promoter by an indirect
 A:Reference number: I48300; MID:9434105; PMID:8065328
 A:Accession: I48300
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-80, 'E', 82-105, 'D', 107-147 <REP>
 A:Cross-References: EMBL:X54876; NID:g50486; PIDN:CAA38657.1; PID:g50487
 C:Genetic:
 A:Gene: COL1A1
 A:Introns: 770/3; 788/3; 806/3; 842/3; 860/3; 878/3; 932/3; 968/3; 1004/3; 1022/3; 1058/
 C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; heterotrimer; triple helix
 F:1-22/Domain: signal sequence #status predicted <Sig>
 F:23-151/Domain: amino-terminal propeptide #status predicted <PRO>
 F:152-1453/Product: von Willebrand factor type C repeat homology <WVC>
 F:1224-1453/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:

Pred. No.:	5,29e-06	Length:	1453
Score:	227.50	Matches:	153
Percent Similarity:	32.05%	Conservative:	30
Best Local Similarity:	26.80%	Mismatches:	230
Query Match:	6.44%	Indels:	159
DB:	2	Gaps:	26

US-10-030-271-3 (1-1883) x S21626 (1-1453)

```

QY 33 CCNAAACAGTGGCTCGCGCCGCGCCAGAGTCATCGGACCCAGAAATCTGCGCGGT 92
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 653 ProGlyAspLeuGlnGlyAlaProGlyProSerGlyAlaArgGlyGluThrGlyPheProGly 672

QY 93 TCTGACTGTTTCCGCTCC-----TCCTCC 119
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 673 GluArgGlyValGlnGlyProGlyProAlaGlyProArgGlyAsnAsnGlyAlaPro 692

QY 120 GGGAAATGGCGCTATCCGGGTCGACCCCGGCC-----GTGCTGGAGAGAGATGAGT 173
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 693 GlyAsnAspGlyAlaAsnGlyAspThrGlyAlaProGlyAlaProGlySerGlnGlyAla 712

QY 174 CCGGACTACTA-----CGGATGCTGTGCTTACCGTATGTTGAGGT 218
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 713 ProGlyLeuGlnGlyMetProGlyGluArgGlyAlaAlaGlyLeuProGlyProGlySerGly 732

QY 219 GGTGGCGGGCAACTGACCGAGTGCAGCTGAGCTCTGCTTCTGCTGATGAGC 278
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 733 -----AspArgGlyAspAlaGlyProLysGlyAlaAspGlySerProGly 747

QY 279 TCTTGCGCGCGCGGAGCTTACCCGCGCGCCAGCGCTGAGACTCTGCTGAGGT 338
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 748 LysAspGlyAlaArgGlyLeuThr---GlyProIleGlyProProGlyProAlaGlyAla 766

QY 339 GGAGCCCGCGG-----GCAGTGGCGGAGAACCACTGCGCTGCGGCAACTCCT 392
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 767 ProGlyAspLysGlyGluAlaGlyProSerGlyProProGlyProThrGlyAlaArgGly 786

QY 393 GCG-----CGTCTGCGCGCGCAAGCTGCT----- 419
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 787 AlaProGlyAspArgGlyGluAlaGlyProProGlyProAlaGlyPheAlaGlyProPro 806

QY 420 -----GCCGCACTGCGCGCGCAAGCGCGCGCGCGAGTCTTCACGAAGCGTATAG 470
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

Db 807 GlyAlaAspGlyGlnProGlyAlaLysGlyGluProGlyAspThrGly-ValLysGlyAs 826
QY 471 CTATGGCACTTCAGCTCTT----- 490
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 826 PAlaGly- ProProGlyProAlaGlyProAlaGlyProProGlyProIleGlyAsnValG 846
QY 491 -----CAAGAGACAGAGGTACCTGCGCGCTCGCGAGTCAAGCACTTC 539
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 846 LysAlaProGlyProLysGlyProArgGlyAlaAlaGlyProProGlyAlaThrGly- 864
QY 540 TCCAAATTCAGCAGCGGTGAGGAGAGCAGAGCTCCCGCCCAACCAAGCGGAGCGG 599
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 865 -----PheProGlyAlaAlaGlyArgValGlyProProGlyProSerGlyAsnAlaG 882
QY 600 GAGTCGGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGCGCCAG----- 652
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 882 LysProProGlyProProGlyProValGlyGlyGlyGlyGlySerGlyProArgGlyGlu 902
QY 653 -----CCGCACTCCAGACAGA 668
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 902 hncLysProAlaGlyAlaArgProGlyGluValGlyProProGlyProProGlyProAlaGlyG 922
QY 669 GTCAGAGCCCGCAGACCTTCTCGAAGGCAAGTGAACCTGACATCCGCGC---TCCG 725
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 922 LysLysGlySerProGlyAlaAspGlyProAlaGlySerProGlyThrProGlyProGlnG 942
QY 726 GGTTCGAGCAGAGTACTGCGACATGCGCAGCTTGA-----GCAGGCGGTGCGATC 779
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 942 LysLysLysGlyGlnArgGlyValGlyLysProGlyGlnArgGlyGluThrGlyPhe 961
QY 780 CCGCGCGCGCCAGCGCGCTGCGCGCAGCTGACCGTGTGG-----GCAGGC 827
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 962 ProGlyLeuProGlyProSerGlyGluProGlyLysGlnGlyProSerGlySerGly 981
QY 828 CACCGCAGT-----GTCGCGTCAAGGAGCACTGGGCTGTGTTGTGACATCAAGTT 881
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 982 GluArgGlyProProGlyProMetGlyProProGlyLeuAlaGlyProProGlyGlu- 1000
QY 882 CTCAGACTCTCTATCTTGACGC----- 905
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1001 -----SerGlyArgGlnGlySerProGlyAlaGlyGlySerProGlyArg 1015
QY 906 -----CTTGGCGCGACTGATGCGCGCTCT----- 938
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1016 AspGlyAlaProGlyAlaLysGlyAspArgGlyGlnThrGlyProAlaGlyProProGly 1035
QY 939 -----GCAGCGCTGCGCGCGCTTCTGACTGA 968
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1036 AlaProGlyAlaProGlyAlaProGlyProValGlyProAlaGlyLysAsnGlyAspArg 1055
QY 969 GGCCTGCGAGAGGCTGTGGGCGCGGAGC---TGTTCGCTGCTGTGTCAGTGTGAGTGA 1025
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1056 GlyLysLysGlyProAlaGlyProIleGlyProAlaGlyAlaArgGly- 1074
QY 1026 GGTGACTATGAGGCTGGCGCGCGCTGTGCTGATGAGAGAGAGGCGGCGG 1085
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1075 -----ProAlaGlyProGlnGlyProArgGlyAspLysGlyGlnThrGlyGlnGln 1091
QY 1086 CCGCAGAGAGGCTCTGATGAGCAGACTGCGAGATGATCCACATCCCAATCTCCGCG 1145
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1092 GlyAspArgGly-----LysGlyLysArgGlyPheSerGlyLeuGln- 1107
QY 1146 CCACT---TCTCTGAGAGAGACACATCTACCTCTG- 1183
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1108 ProProGlySerProGlySerProGlyGlnGlnGlyProSerGlyAlaSerGlyProAla 1127
QY 1184 -----ACAGCCCTCCACAGAGATGGCTCTAGGCTTAACATTTCCAGCTGAGT 1237
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1128 GlyProArgGlyProProGlySerAlaGlySerProGlyLysAsp-----GlyLeuAsn 1145
QY 1238 TTCCTTCCAGAGCTCTCTTACCCCGAGGTG- 1270
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1146 GlyLeuProGlyProIleGlyProProGlyProAlaGlyArgThrGlyAspSerGlyPro 1165

```

QY 1271 -----CCCTAGCTCGAGGCGGGGCTGGCTGTATCTCAGAGGAGGCA 1324
|||
DB 1166 AAGlyProProGlyProProGlyProProGlyProProSerGlyGlyTyr 1185
|||
QY 1325 CAGCTACACACTCACCAGGCCCC 1351
|||
DB 1186 AspPheSerPheLeuProGlnProPro 1194
|||
RESULT 14
CGCH15
collagen alpha 1(I) chain - chicken (tentative sequence) (fragments)
C/Species: Gallus gallus (chicken)
C/Date: 12-Aug-1981 #sequence, revision 06-Jul-1982 #text_change 31-Mar-2000
C/Accession: A90458; A90181; A02857
R/Higberger, J.H.; Corbett, C.; Dixit, S.N.; Yu, W.; Seyer, J.M.; Kang, A.H.; Gross, J.
Biochemistry 21, 2048-2055, 1982
A/File: Amino acid sequence of chick skin collagen alpha1(I)-C88 and the complete prime
A/Reference number: A90458; MUID:8231995; PMID:7093229
A/Accession: A90458
A/Molecule type: protein
A/Residues: 1-1036 <HIG>
A/Experimental source: skin
A/Note: This is the latest in a series of papers from these workers elucidating the sequ
R/Byre, D.R.; Glimcher, M.J.
Biochem. Biophys. Res. Commun. 48, 720-726, 1972
A/File: Evidence for a previously undetected sequence at the carboxyterminus of the alp
A/Reference number: A90181; MUID:72243016; PMID:5047697
A/Accession: A90181
A/Molecule type: protein
A/Residues: 1037-1042 <BYR>
A/Experimental source: skin
A/Note: residues 1037-1042 above correspond to the carboxyl end of the protein
C/Comment: Lysines at positions 103, 700, 934, and 946 above may be hydroxylated in some
C/Comment: Pro-1002 is the only 3-hydroxyproline and the only hydroxylated proline in po
C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C/Keywords: cold coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer;
F./Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Alignment Scores:
Pred. No.: 7.58e-06 Length: 1042
Score: 225.00 Matches: 175
Percent Similarity: 31.23% Conservative: 33
Best Local Similarity: 26.28% Mismatches: 228
Query Match: 6.56% Indels: 230
DB: 1 Gaps: 36

US-10-030-271-3 (1-1883) x CGCH15 (1-1042)

QY 1675 TCTGACAGCAGGGG-----GGAAGCTGAGAGAGAGTCAAGTCTGTAAATT 1631
|||
DB 256 SerGlyProGlnGlyProSerGlyAlaProGlyProGlySerGlySerGly----- 272
|||
QY 1630 GAAGCTGGGAGGAGCCCTGCTGGCTGGAGATGTGTGGCAAG-----TGAG 1583
|||
DB 273 GluProGly--AlaProGlyAsnLysGlyAspThrGlyAlaLysGlyGluProGlyPro 232
|||
QY 1582 CAGGCCCATGTGACACCCGAGCTCAAGTCCACTAT-----GTGGCTGAAGCCC 1532
|||
DB 292 IaglyValGlnGlyProProGlyProAlaGlyGlnGlyLysGlyValAlaArglyG 312
|||
QY 1531 AGGTGTGTGTGTCATTAAGAGTCTGTAGAGAGAGAGCCCTTAGGCCAAAGAGTGTG 1472
|||
DB 312 IuProGlyProAlaGlyLeuProGlyProAlaGlyGlnArglyAlaPro----- 328
|||
QY 1471 TGCTGGAGACAG-----CCAGGCGAGAGGCCA 1442
|||
DB 329 -----GlySerArglyPheProGlyAlaAspGlyIleAlaGlyProGlyGlyProProG 347
|||
QY 1441 GCGAAGATCTCAGCAGGCGCAGTGTGACATGACGAGTGTGTGCACCTGTGACATGTCA 1382
|||
DB 347 IyGlnArglySerProGlyAlaValGly----- 356
|||

QY 1381 GACAGCCCAAGATCAGATATCAATGTGCAAGGGGAGCTTTGTGAGTGT----- 1332
|||
DB 357 -----Prolys-----GlySerProGlyGlnAlaGlyAArgProG 368
|||
QY 1331 -----GTAGCTGTCCCTCTCTCTTGTGAGATACAGGCCACGCC-----CGCCTCTCCG 1283
|||
DB 368 IyGlnAlaGlyLeuProGlyAlaLysGlyIleThrGlySerProGlySerProGlyPro 388
|||
QY 1282 AGGCTAAGGGGGGACACCTGGGGGTGAGAGAGTCTGGGAAGAACTCAGCTGGAATTG 1223
|||
DB 388 spGlyIySerThrGlyProProGly----- 395
|||
QY 1222 GTTTAGGCTCAGAGCCCATCTGTGGAGGGGCTGTCAAGGGGTAGAGATGTCTGTC 1163
|||
DB 396 -----ProAlaGlyGlnAspGlyAArgP 403
|||
QY 1162 CTCCAGAGAGAGTGTGCGGAGACTGTGAGTGTGATCAATCTGCCAGTCTTGATC 1103
|||
DB 403 roGlyProAlaGlyProProGlyAlaArgGlyGlnAlaGlyValMetGlyPheProGlyP 423
|||
QY 1102 AGGAGGCTCTGTGCGG-----CGCGCCCTCTCTCTCTCA 1064
|||
DB 423 rolyserGlyAlaAlaGlyGluProGlyLysProGlyGlnArglyAlaProGly-ProPro 442
|||
QY 1063 TCAGCAACAGGCGCGCGCGCGAG-----CCTCATGTACAGCTCATCA 1019
|||
DB 443 GlyAlaValGlyAlaAlaGlyLysAspGlyAlaGlyAlaGlnGlyProProGlyPro 462
|||
QY 1018 CACTGACCAAGCAGCAGCAAG-----CTCCCGGCCCAAGCTCTCCAGGCGCTCATCA 962
|||
DB 463 ThrGlyProAlaGlyIuArglyGlnGlnGlyProAlaGlyAlaProGlyPheGlnGly 482
|||
QY 961 GGAACAGCGCCGCGAGGCT-----GCAAGGAGCGCG----- 929
|||
DB 483 LeuProGlyProAlaGlyProProGlyGlnAlaGlyLysProGlyGlnGlnGlyAlPro 502
|||
QY 928 -----CACTCAGTGTGCGGCCCAAGAGGCT--CCAGATAGG 893
|||
DB 503 GlyAsnAlaGlyAlaProGlyProAlaGlyAlaArgGlyGlnArglyPheProGlyGln 522
|||
QY 892 AGAGCTGTGAACTTGATGTCAAAACCAAGAGCCCAAGTCTTGAGCGCAGCATG 833
|||
DB 523 ArgGlyValGlnGly--ProProGlyProGlnGlyProArgGlyAlaAsnGlyAlaPro 541
|||
QY 832 CGGTGCGCTGCCAAACAGCTCAGCTGCCCGCCACAG----- 794
|||
DB 542 GlyAsnAspGlyAlaLysGlyAspAlaGlyAlaProGlyAlaProGlyAsnGlnGlyPro 561
|||
QY 793 CTGGGGCGCGCGGATGCCA-----CGCCTGTCTCAAGGCTGGCCCATGTCTGCAGT 740
|||
DB 562 ProGlyLeuGlnGlyMetProGlyGlnArglyAlaAlaGlyLeuProGlyAlaLysGly 581
|||
QY 739 ACTCTCTCGAACCGGAGCG-----GATGTCAAGTCACTTGTCTTCAG 692
|||
DB 582 AspArgGlyAspProGly-ProLysGlyAlaAspGlyAlaProGlyLysAspGlyLeuArg 601
|||
QY 691 AGGAAGTCTGCGGGGCTCACTGCTGCTGGGTGCGGTGGGGCCCTCTCCGCCGCC 632
|||
DB 601 gGlyLeuThrGlyProIle-----GlyPro--ProGlyPr 612
|||
QY 631 GTCTGCACC-----ACCACTGGCGCGGCC----- 607
|||
DB 612 oAlaGlyAlaProGlyAspLysGlyGlnAlaGlyProProGlyProAlaGlyProThrG 632
|||
QY 606 -----CCGACTCGCGCGCTCCCGCTGTGGGG 578
|||
DB 632 yAlaArgGlyAlaProGlyAspArgGlyGlnProGlyProProGlyProAlaGlyIleAl 652
|||
QY 577 GCGAGCTGTCTCCCACTGAC-----CTGCTGAAGATTTGGAAGAACTGTTGACT 527
|||
DB 652 agly-ProProGly-AlaAspGlyGlnProGlyAlaLysGlyGlnThrGlyAspAlaGly 671
|||


```

Db      686  AAlaGlyProAlaGlyGlySerGlyProAlaGlyProProGlyProGln----- 701
QY      1268  CACCTGGGGGTAGAGAAAGTCTGGAAAGAACTCAGCTGAATGTTTAGGCTCAGA 1209
Db      702  -----GlyValIleGly----- 705
QY      1208  GCCACATCCTGTGGAGGGGCTGTCAAGGGGTAGAGATGATGCTCTCCAGAGAAAG 1149
Db      706  -----GluArgGlySerProGlyGlyProGlyAlaAlaGly 717
QY      1148  TGCCCGGAGACTTGGAGTGGATCAATCTGCAGTCTGATCAGAGAGCCTCTGTC 1089
Db      718  PheProGly----- 720
QY      1088  GGGCGCGCGCCCTTCTCTCTCATATCAGCAACAGGGCGCGCCAGCTCATACTCA 1029
Db      721  GLyArgGlyProProGly-ProProGlySerAsnGlyAsnProGlyProProGlySerSe 740
QY      1028  GCCTCATCCACACTGACACAGGAGCAACAGCTCCGCGCCCAAGCCTTCGAGAGGCC 969
Db      740  rGlyAla-----ProGlyIleAspGlyProProGlyPro-----Pr 752
QY      968  TCAGTCAGAAACAGCGCCCGCAGGGCTGCA--GCAGGGCGCCACTCAGTAGTGGCC 912
Db      752  oGlySerAsnGlyAlaProGlySerProGlyIleSerGlyProGlyGlyAspSerGlyPr 772
QY      911  CAGAAGCGCTCCAGATAGAGAGCTGAGAACTGATGTCAAAACACAGAGCCGAGG 852
Db      772  o-----ProGlyGlyIleArgGlyAlaProGlyProGlyProGlyAlaProGly 789
QY      851  TCCCTTAGCGGACACTGCGGTGCTGCTGCCAAACAGTCCAGCTCCGCGCCAGCGCC 792
Db      789  yProLeuGlyIleAlaGlyLeuThrGlyAlaArgGlyLeuAlaGlyProProGlyMetPr 809
QY      791  TGGGGCGCGCGGATGCCAGCCCTGC-----TCAAAGCTGGCCC 751
Db      809  oGlyAlaArgGlySerProGlyProGlyIleGlyGlyIleGlyAsnGlyIleAspProGlyPr 829
QY      750  ATGCTGCACTACTCTGCTGA-----ACCGGAGCCGGATGTC 712
Db      829  oSerGlyIleAsnGlyGlyIleArgGlyProProGlyProGlyIleuProGlyLeuAlaGly 849
QY      711  ACAGTCACTTGTGCTCAGAGGAAGTCTGGCGGCTGACTGCTGGGGTGGCGGC 652
Db      849  yThrAlaGlyIleuProGlyIleArgGlyAsnProGlySerAspGlyLeuProGlyIleArgAs 869
QY      651  TGGGGCGCTCTCCGCGCGCTGAGCAACCACTGAGCGCGCCGACTCCGCGCTG 592
Db      869  rGlyAlaPro-GlyAlaIleGlyIleAspArgGlyIleAsnGlySerProGlyAlaProGlyAla 889
QY      591  CCGCTTGTTGGGGGAGCTGTCTCCACTGACCTGTGAGAAATTGGCAAACTGCT 532
Db      889  laProGlyIleProGlyProProGlyProAlaGlyProAlaGlyIle----- 904
QY      531  TGAATGCGAGCGAGGAGCTAACCTCTGCTCTTTGAAGAGCTGAGAGTGCATA 472
Db      905  --SerGlyAspArgGlyIleThrGlyPro----- 913
QY      471  GCTATAGCGTTCTGAGACACTGGCGCGCGCTTGGCGCGCGAGTGCAGAGATC 412
Db      914  -----AlaGlyProSerGlyAlaProGlyProAlaGlySerAla 926
QY      411  GTGGCGGCGCAGACGCGCAGAGT----- 386
Db      926  rGlyProProGlyProGlyIleProArgGlyIleAspGlyIleuThrGlyIleuArgGlyAla 946
QY      385  -----GCCCGAGCGCGCAG 370
Db      946  laMetGlyIleuGlyIleuArgGlyIleuProGlyIleuProGlyAlaProGlyIleuSerProG 966
QY      369  GT--TGCTCTCGCGGACTGCGCGCGCTCCAGCTCCAGCAGAGCTCTAGCGCGCT 313

```

```

Db      966  lProAlaGlyIleuGlyIleuAlaValaGlySerProGlyProAlaGlyProArgGlyPro 986
QY      312  GCGGCGCGCGGCTCAAGCTCCGCGCGCGGAGGCTCATCCAGCAAGAAAGCCAGAG 253
Db      986  aGlyProSer-----GlyProProGlyIleuAspGlyAla 997
QY      252  CTCAGCTGCACTCGGTGAGTGGCCGCGCACCTCGAATACGATGGAAGCGACAG 193
Db      997  laSerGlyIleProGlyProIleGlyProProGlyProArgGlyIleuArgGlyIleuArg 1017
QY      192  CATCCGTAAGTATCCAGGCACTCATCT-----CTCCAGACAGCGGCGCGGCTCGA 139
Db      1017  lSerGlyIleSerProGly--HisProGlyIleProGlyProProGlyProProGly- 1035
QY      138  CCGGATAGCGCATTCGCG-----GGGAGGAGAGC 107
Db      1036  -----AlaProGlyProGlyCysGlyAlaGlyIle 1045

```

Search completed: March 22, 2005, 01:55:38
 Job time : 111.105 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 01:03:41 ; Search time 53.9542 Seconds
(without alignments)
5210.505 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 3535
Sequence: 1 aggcgaatcatagagaag.....ctgaaaggcataggt999 1883

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p,model -DEV=x1p
-Q=/cgn2_1/USFTO_spool_p/US10030271/runat_21032005_153459_26492/app_query.fasta_1.3150
-DB=Issued_Patents_AA -QFMT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10030271 @CGN 1.1 81 @runat_21032005_153459_26492 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*
 - 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pep:*
 - 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*
 - 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*
 - 5: /cgn2_6/ptodata/1/iaa/PTUS.COMB.pep:*
 - 6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	647.5	18.3	318	3	US-08-859-167-4
2	647.5	18.3	318	3	US-09-109-273-4
3	647.5	18.3	318	3	US-09-276-993-4
4	647.5	18.3	318	4	US-09-723-450-4
5	647.5	18.3	318	4	US-09-733-167A-1
6	641.5	18.1	318	3	US-08-859-167-6
7	641.5	18.1	318	3	US-09-109-273-6
8	641.5	18.1	318	3	US-09-276-993-6
9	641.5	18.1	318	4	US-09-723-450-6
10	641.5	18.1	318	4	US-09-733-167A-3
11	455	12.9	210	4	US-09-733-167A-6
12	286	8.3	1411	4	US-09-252-991A-28408

C 13	268.5	7.8	504	3	US-09-219-849-3	Sequence 3, Appli
C 14	268.5	7.8	561	1	US-08-642-255-52	Sequence 56, Appl
C 15	262.5	7.7	726	1	US-09-252-991A-20675	Sequence 20675, A
C 16	261.5	7.6	957	4	US-09-252-991A-20408	Sequence 20408, A
C 17	261	7.6	720	3	US-09-219-849-4	Sequence 4, Appli
C 18	261	7.6	777	1	US-08-642-255-53	Sequence 53, Appl
C 19	260	7.6	661	4	US-09-252-991A-24567	Sequence 24567, A
C 20	257	7.5	450	4	US-09-252-991A-16659	Sequence 16659, A
C 21	254	7.4	802	4	US-09-252-991A-25050	Sequence 25050, A
C 22	251	7.3	1008	4	US-09-252-991A-29419	Sequence 29419, A
C 23	250	7.1	663	4	US-09-252-991A-30843	Sequence 30843, A
C 24	250	7.1	957	4	US-09-252-991A-20408	Sequence 20408, A
C 25	249.5	7.3	1064	1	US-08-642-255-62	Sequence 62, Appl
C 26	246	7.2	542	4	US-09-252-991A-21753	Sequence 21753, A
C 27	245.5	7.2	1706	4	US-09-252-991A-31263	Sequence 31263, A
C 28	245.5	7.2	2294	4	US-09-252-991A-17231	Sequence 17231, A
C 29	243.5	7.1	663	4	US-09-252-991A-30843	Sequence 30843, A
C 30	243.5	7.1	863	4	US-09-252-991A-26099	Sequence 26099, A
C 31	242.5	7.1	588	4	US-09-252-991A-30690	Sequence 30690, A
C 32	242.5	7.1	1034	4	US-09-252-991A-28921	Sequence 28921, A
C 33	242	7.1	686	4	US-09-252-991A-20509	Sequence 20509, A
C 34	242	7.1	809	4	US-09-252-991A-31759	Sequence 31759, A
C 35	241	7.0	885	4	US-09-252-991A-26129	Sequence 26129, A
C 36	241	7.0	973	4	US-09-252-991A-32186	Sequence 32186, A
C 37	240.5	7.0	622	4	US-09-252-991A-32308	Sequence 32308, A
C 38	240.5	7.0	1418	4	US-09-252-991A-32367	Sequence 32367, A
C 39	239	7.0	369	4	US-09-252-991A-25394	Sequence 25394, A
C 40	238	6.9	664	4	US-09-252-991A-30396	Sequence 30396, A
C 41	237.5	6.9	960	3	US-09-219-849-6	Sequence 6, Appli
C 42	237	6.7	783	4	US-09-252-991A-18035	Sequence 18035, A
C 43	237	6.9	1065	1	US-08-642-255-72	Sequence 72, Appl
C 44	236.5	6.9	491	4	US-09-489-039A-7836	Sequence 7836, Ap
C 45	236.5	6.9	869	4	US-09-252-991A-22290	Sequence 22290, A

ALIGNMENTS

RESULT 1
US-08-859-167-4
Sequence 4, Application US/08859167
Patent No. 6037461
GENERAL INFORMATION:
APPLICANT: Alnemir, Emdad S.
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
TITLE OF INVENTION: OF MAKING THE SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461 is
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,167
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
;

LENGTH: 318 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-859-167-4

Alignment Scores:

Pred. No.:	9,846-39	Length:	318
Score:	647.50	Matches:	141
Percent Similarity:	63.55%	Conservative:	56
Best Local Similarity:	45.48%	Mismatches:	92
Query Match:	18.32%	Indels:	21
DB:	3	Gaps:	6

US-10-030-271-3 (1-1883) x US-08-859-167-4 (1-318)

```

OY 157 TGGAGAGAGATAGATGCTGAGCTACTACGGAGTCTGCTGCTTACCGTATGTTGAG 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12 TTPProgluGlnHisglYgluGlnGlnHisglYleuTYrSerleuHisArgmetPheAsp 31
OY 217 GTGGTGGGCGGCGCACTGACCGAGCTGAGCTGCTGCTCTTCTG--CTGAT 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 32 ILevalGlyThrHisleuThrHisArgAspValArgValleuSerPheleuPheValAsp 51
OY 274 GAGGCTCTGGCGCGCGCGAGCTTAAGCCGGCGCGCGAGCGGCTTAAGCTCTGCTG 333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 52 ValIleAspAspHisglYArgglYleu-----IleArgAsnGlyArgAspPheleu 69
OY 334 GAGCTGAGCGCGCGCGCGAGCTGCGCGAGAGCACTGCGCTGCTGGGGCAACTCTCTG 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 70 AlaleuGlnArgGlnIlyArgCysAspGlnSerAspPheArgGlnValleuGlnleu 89
OY 394 CGCGTGTGCGCGCGCGAGCACTGCTGCTGCGCACTGCGCGAGCGCGCGCGCGCGAGTG 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 90 ArgIleIleThrArgHisAspLeuLeuProTYrValThrleuYsArgArgalaVal 109
OY 454 TCTCCA-----GAGCGTCTAGCTATGAGCACTCC-----AGCTCTTAAAG 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 110 CysProAspLeuValAspYsTYrleuGlnGlnThrSerIleArgTYrValThrProArg 129
OY 496 AGGACAGAGAGTATGCTGCGCGCTGCGAGTACAGCAAGTCTGCAAAATCTCAGCAG 555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 130 AlaleuSerAspProgluProArgProProGlnProSerIlyThrValProProHISTYr 149
OY 556 GGTCAG-----TGGGAGACAGGCTCCCGCCCAACCAAGCGAGCGGCGAGCTCGGGC 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 150 ProValValCysCysProThrSerGlyProGlnMetCysSerIysArgProAlaArgIly 169
OY 610 CGGCGCAAGTGTGTGTCAGACGCGCGCGAGAGAGGGGCGCGCGCGCGCGAGCAGCAG 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 170 ArgAlaThrleuGlnYserGlnArgIysArgArg----- 180
OY 670 TCAGAGCCCGCGAGACTTCTCTCTGTAAGGCAAGTACCTGATGATCCGGGCTT 729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 ---LysSerValThrProAspProIySglnYsGlnThrCysAspIleArgleuArgVal 199
OY 730 CGAGCAGACTACTGCGAGCAGATGGCGCGCTTGAGCAGAGCGCGCTGACCTCCGCGCGCC 789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 200 ArgAlaGlnIlyCysGlnHisglYThrAlaIleuGlnGlyAsnValPheSerAsnIySgln 219
OY 790 CAGGCGCTGCGCGCGAGCTGACGCTGTTGGGAGCGCAACCGCACTGCTGCGCTCAAG 849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 220 AspProleuGlnArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleleuYsSerArg 239
OY 850 GACCTGGGCTGTGTGTGTGATCATCAAGTCTCAGAGCTCTATATGAGCGGCTTC 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 240 AspleuGlySerIleIleCysAspIleYsPheSerGlnleuThrYrleuAspAlaPhe 259
OY 910 TGGGGGACTACTGCTGAGCGCGCGCTGCTGAGCGCTTCGAGGCGCTGTTCTGACTAG 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 260 TPrArgAspTYrIleAsnGlySerleuGlnValIleuYsGlyAlaPheIleThrAsp 279
OY 970 GCCCTGCGAGAGGCTGTGGCGCGGAGGCTGTTCGCTGCTGCTGATGATGAGGCT 1029
  
```

```

DB 280 SerleuYsGlnAlaValaIleYHisglYalIleYsleuValaAsnValAspGlnGlu 299
OY 1030 GACTATGAGCTGCGCGCGCGCGCTGTTG 1059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 300 AspTYrGlnleuGlyArgGlnIySleuLeu 309
  
```

RESULT 2

US-09-109-273-4
 Sequence 4, Application US/09109273
 Patent No. 6063760

GENERAL INFORMATION:

APPLICANT: Alnemi, Emad S.
 APPLICANT: Fernandez-Alnemi, Teresa
 TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
 TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
 NUMBER OF INVENTION: OF MAKING THE SAME
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760rIs
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: WINDOWS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/109,273

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/859,167

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 318 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-109-273-4

Alignment Scores:

Pred. No.:	9,846-39	Length:	318
Score:	647.50	Matches:	141
Percent Similarity:	63.55%	Conservative:	56
Best Local Similarity:	45.48%	Mismatches:	92
Query Match:	18.32%	Indels:	21
DB:	3	Gaps:	6

US-10-030-271-3 (1-1883) x US-09-109-273-4 (1-318)

```

OY 157 TGGAGAGAGATAGATGCTGAGCTACTACGGAGTCTGCTGCTTACCGTATGTTGAG 216
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12 TTPProgluGlnHisglYgluGlnGlnHisglYleuTYrSerleuHisArgmetPheAsp 31
OY 217 GTGGTGGGCGGCGCACTGACCGAGCTGAGCTGCTGCTCTTCTG--CTGAT 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 32 ILevalGlyThrHisleuThrHisArgAspValArgValleuSerPheleuPheValAsp 51
OY 274 GAGGCTCTGGCGCGCGCGAGCTTAAGCCGGCGCGCGAGCGGCTTAAGCTCTGCTG 333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 52 ValIleAspAspHisglYArgglYleu-----IleArgAsnGlyArgAspPheleu 69
  
```

```

QY 334 GAGCTGAGCGCCGCGGCAAGTGGCGGAGCAACCTGCGGCTGCGGCAATCTCTG 393
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 70 AlaleuGluArgInglInglArgCyAspGluSerAsnPhaArgInValLeuGlnLeu 89
QY 394 CGCGTGTGCGCCCGCAAGCTGCTGCGCACTGGCGCGCAAGCGCGCGCGCACTG 453
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 90 ArgIleIleThrArgHisAspLeuLeuProTyrValThrLeuValSarGArgAlaVal 109
QY 454 TCTCCA-----GAACGCTATAGCATGGCACTCC-----AGCTTTCAAG 495
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 110 CyAspProAspLeuValAspLysTyrLeuGluGlnThrSerIleArgTyrValThrProArg 129
QY 496 AGGACAGAGGTAAGCTCCGCGCGCTGCGGAGTCAAGCAAGTTCTGCAATTCTCAGCAG 555
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 130 AlaleuSerAspProGluProArgProGluInProSerLysThrValProProHisIleTyr 149
QY 556 GGTTCAG-----TGGAGACAGGCTCCCGCCCAACCAAGCGCGCGGAGTGGGCGC 609
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 150 ProValValCyAspProThrSerGlyProGluMetCySerLysArgProAlaArgGly 169
QY 610 CGGCGCAAGTGTGTGTCAGCGCGCGCGGAGAGGCGCGCCAGCCGACCCAGCAGCAG 669
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 170 ArgAlaThrLeuGlySerGlnArgLysArgArg----- 180
QY 670 TCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAAGTGAATCGGCTCGGAGTT 729
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 181 ---LysSerValThrProAspProLysGlnLysGlnThrCyAspAlaIleArgLeuArgVal 199
QY 730 CGAGCAGAGTACTGCGAGCATGGCGCAAGCTTGAAGCAGGCGGTGGCATCCCGCGCGCC 789
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 200 ArgAlaGluTyrCySglnHisGluThrAlaLeuGlnGlnLysAsnValPheSerAsnLysGln 219
QY 790 CAGCGCTGCGCGCGCGCAAGTGAAGTGTGGAGAGCGCAAGCTGCTGCTGCAAG 849
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 220 AspProLeuGluArgGlnPheGluArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
QY 850 GACCTGGAGCTGTGCTTGTGATCATCAAGTTCTCAGAGCTCTATCTGAGAGCTTC 909
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 240 AspLeuGlySerIleIleCyAspAlaIleLysPheSerGlnLysThrTyrLeuAspAlaPhe 259
QY 910 TGGGGCAGTACTGAGTGGCGCGCTGCTGTCAGAGCGCTGCGGCGCGTGTTCAGTGA 969
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 260 TrpArgAspTyrIleAsnGlySerLeuGlnLysValLeuLysGlyValPheIleThrAsp 279
QY 970 GCCCTGGAGAGCTGTGGCGCGGAGGCTGTGCGCTGCTGTCAGTGTGAAGAGCT 1029
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGluGln 299
QY 1030 GACTATGAGCTGGCGCGCGCGCGCTGTG 1059
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 300 AspTyrGlnLeuGlnArgGlnLysLeuLeu 309

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,993
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-276-993-4

Alignment Scores:
Pred. No.: 9,84e-39 Length: 318
Score: 647.50 Matches: 141
Percent Similarity: 63.55% Conservative: 56
Best Local Similarity: 45.48% Mismatches: 92
Query Match: 18.32% Indels: 21
DB: 3 Gaps: 6

US-10-030-271-3 (1-1883) x US-09-276-993-4 (1-318)
QY 157 TGGAGAGAGATGAGTGCCTGAGTACTACGGAGTGTGCTTCAACCTGATTCGAG 216
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 12 TrpProGluGlnHisGlyGlnGlnGlnHisGlyLeuTyrSerLeuHisSarGMetPheAsp 31
QY 217 GTGTGGGCGCGCACTGACCGAGTGGAGTGGAGTGGAGTCTGCGGCTTCG---CTGAT 273
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 32 IleValIleLysThrHisLeuThrHisSarGAspAlaArgValLeuSerPheLeuPheValAsp 51
QY 274 GAGGCTCTGAGCGCGCGCGGAGGCTTACCGCGCGCGCAAGCGCTGAGCTCTGCTG 333
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 52 ValIleAspAsnHisGluArgGlyLeu-----IleArgAsnGlyAlaGAspPheLeu 69
QY 334 GAGCTGAGCGCGCGCGGCAAGTGGCGGAGCAACCTGCGGCTGCGGCAATCTCTG 393
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 70 AlaleuGluArgInglInglArgCyAspGluSerAsnPhaArgInValLeuGlnLeu 89
QY 394 CGCGTGTGCGCCCGCAAGCTGCTGCGCACTGGCGCGCAAGCGCGCGCGCACTG 453
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 90 ArgIleIleThrArgHisAspLeuLeuProTyrValThrLeuValSarGArgAlaVal 109
QY 454 TCTCCA-----GAACGCTATAGCATGGCACTCC-----AGCTTTCAAG 495
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 110 CyAspProAspLeuValAspLysTyrLeuGluGlnThrSerIleArgTyrValThrProArg 129
QY 496 AGGACAGAGGTAAGCTCCGCGCGCTGCGGAGTCAAGCAAGTTCTGCAATTCTCAGCAG 555
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 130 AlaleuSerAspProGluProArgProGluInProSerLysThrValProProHisIleTyr 149
QY 556 GGTTCAG-----TGGAGACAGGCTCCCGCCCAACCAAGCGCGCGGAGTGGGCGC 609
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 150 ProValValCyAspProThrSerGlyProGluMetCySerLysArgProAlaArgGly 169
QY 610 CGGCGCAAGTGTGTGTCAGCGCGCGCGGAGAGGCGCGCCAGCCGACCCAGCAGCAG 669
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 170 ArgAlaThrLeuGlySerGlnArgLysArg----- 180
QY 670 TCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAAGTGAATCGGCTCGGAGTT 729
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

```

Db 181 ---LysSerValThrProAspProLysGluLysGlnThrCysAspIleArgLeuArgVal 199
 QY 730 CGAGCAGAGTACTGCGAGCATGGGCGACCTTGGAGCAGGCGCTGGACCTCCGCGCGCC 789
 Db 200 ArgAlaGluTyrCysGlnHisGluThrAlaLeuGlnGlyAsnValPheSerAsnLysGln 219
 QY 790 CAGGCGCTGGCGGCGAGCTGGACGCTTTGGGCGAGCCAGCCGACATGCTGGCGCTCAAG 849
 Db 220 AspProLeuGluArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
 QY 850 GACCTGGGCTCTGTGTTGTGACATCAATTTCTGAGGCTCTCTATCTGAGCGCTTC 909
 Db 240 AspLeuGlySerIleIleCysAspIleLysPheSerGlnLeuThrTyrIleAspAlaPhe 259
 QY 910 TGGGCGGACTACCTGAGTGGCGCGCTGCTGAGCGCTCGCGGCGCTTCTCTGACTGAG 969
 Db 260 TTPArgAspTyrIleAsnGlySerLeuLeuGluAlaLeuLysGlyValPheIleThrAsp 279
 QY 970 GCCCTGCGAGAGCTGTGGCGCGGAGGCTGTTCGCTGTGCTGCTGCTGATGAGGCT 1029
 Db 280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
 QY 1030 GACTATGAGGCTGGCGGCGCGCTGTGG 1059
 Db 300 AspTyrGlnLeuGlyArgGlnLysLeuLeu 309

RESULT 4
 US-09-723-450-4
 ; Sequence 4, Application US/09723450
 ; Patent No. 6576751
 ; GENERAL INFORMATION:
 ; APPLICANT: Alnemri, Emad S.
 ; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, An
 ; FILE REFERENCE: TJU2445
 ; CURRENT APPLICATION NUMBER: US/09/723,450
 ; CURRENT FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 09/276,993
 ; PRIOR FILING DATE: 1999-03-26
 ; PRIOR APPLICATION NUMBER: 08/859,167
 ; PRIOR FILING DATE: 1997-05-20
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 318
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: No. 6576751el Sequence
 ; US-09-723-450-4

Alignment Scores:
 Pred. No.: 9 84e-39 Length: 318
 Score: 647.50 Matches: 141
 Percent Similarity: 63.55% Conservative: 56
 Best Local Similarity: 45.48% Mismatches: 92
 Query Match: 18.32% Indels: 21
 DB: 4 Gaps: 6

US-10-030-271-3 (1-1883) x US-09-723-450-4 (1-318)

QY 157 TGGGAGAGAGATAGTGCCTGACTACAGGAGTCTGTGCTGACCGTATGTTGAG 216
 Db 12 TTPProGlnGlnHisGlyGlnGlnHisGlyLeuTyrSerLeuHisArgMetPheAsp 31
 QY 217 GTGGTGGGCGGCGAATGACCGAGTGGAGCTGAGCTCTGCGCTTTCTG---CTGAT 273
 Db 32 lIeValGlyThrHisLeuThrHisArgAspValArgValLeuSerPheLeuPheValAsp 51
 QY 274 GAGGCTCTGGCGCGCGCGAGGCTTAGGCGCGCGCGCGAGGCGCTTAGAGCTCCGCTG 333
 Db 52 ValIleAspAspHisGlnGluArgGlyLeu-----lIeArgAsnGlyArgAspPheLeuLeu 69

QY 334 GACTGGAGCGCCGCGGCGAGTGGCGCGAGAGCAACTGCGGCTGCTGCGGCGAATCTCTG 393
 Db 70 AlaLeuGluArgGlnGlyArgCysAspGlnSerAsnPheArgGlnValLeuGlnLeuLeu 89
 QY 394 CGCGTGGCGCGCGCGAGCACTGCTGCCGCGACCTGGCGCGGAGCGGCGCGCGCGAGT 453
 Db 90 ArgIleIleThrArgHisAspLeuLeuProTyrValThrLeuLysArgArgAlaVal 109
 QY 454 TCTCCA-----GAACGTATAGCTATGACACTCC-----AGCTTCTCAAG 495
 Db 110 CysProAspLeuValAspLysTyrLeuGlnGlnThrSerIleArgTyrValThrProArg 129
 QY 496 AGGACAGAGGTAAGCTGCGCGTCCGCGAGTCAAGCAATTCGAAATTCAGCAG 555
 Db 130 AlaLeuSerAspProGlnProArgProGlnProSerLysThrValProProHisTyr 149
 QY 556 GGTGAG-----TGGGAGACAGGCTCCCGCCCAACCAAGCGGACGCGCGCGCGGCG 609
 Db 150 ProValValCysCysProThrSerGlyProGlnMetCysSerLysArgProAlaArgGly 169
 QY 610 CGGCCAGTGTGTGTGCCAGAGCGGCGGAGAGGAGGCGCCAGCGCAAGCGAGCAG 669
 Db 170 ArgAlaThrLeuGlySerGlnArgLysArgArg----- 180
 QY 670 TCAGAGCGCCGACACTCTCTGTGAAGCAAGTGAAGCTGTGACATCCGCTCGGCTT 729
 Db 181 ---LysSerValThrProAspProLysGlnLysGlnThrCysAspIleArgLeuArgVal 199
 QY 730 CGAGCAGAGTACTGCGAGCATGGGCGACCTTGGAGCAGGCGGCGCATCCGCGCGCC 789
 Db 200 ArgAlaGluTyrCysGlnHisGluThrAlaLeuGlnGlyAsnValPheSerAsnLysGln 219
 QY 790 CAGGCGCTGGCGGCGAGCTGGACGCTTTGGGCGAGCCAGCCGACATGCTGGCGCTCAAG 849
 Db 220 AspProLeuGluArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
 QY 850 GACCTGGGCTCTGTGTTGTGACATCAATTTCTGAGGCTCTCTATCTGAGCGCTTC 909
 Db 240 AspLeuGlySerIleIleCysAspIleLysPheSerGlnLeuThrTyrIleAspAlaPhe 259
 QY 910 TGGGCGGACTACCTGAGTGGCGCGCTGCTGAGCGCTCGCGGCGCTTCTCTGACTGAG 969
 Db 260 TTPArgAspTyrIleAsnGlySerLeuLeuGluAlaLeuLysGlyValPheIleThrAsp 279
 QY 970 GCCCTGCGAGAGCTGTGGCGCGGAGGCTGTTCGCTGTGCTGCTGATGAGGCT 1029
 Db 280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
 QY 1030 GACTATGAGGCTGGCGGCGCGCTGTGG 1059
 Db 300 AspTyrGlnLeuGlyArgGlnLysLeuLeu 309

RESULT 5
 US-09-733-167A-1
 ; Sequence 1, Application US/09733167A
 ; Patent No. 6696547
 ; GENERAL INFORMATION:
 ; APPLICANT: Kramer, Peter
 ; TITLE OF INVENTION: Protein for Regulation of Apoptosis
 ; FILE REFERENCE: 4121-120
 ; CURRENT APPLICATION NUMBER: US/09/733,167A
 ; PRIOR APPLICATION NUMBER: 2000-12-08
 ; PRIOR FILING DATE: 1999-06-08
 ; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
 ; PRIOR FILING DATE: 1998-06-08
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 318
 ; TYPE: PRT


```

Db      90  ArgillelThrArgHisAspLeuProTyValThrLeuLysAlaVal 109
Qy      454  TCTCCA-----GACGCTATAGCTATGGCACTCCAGCTTCAAGAGACAG 504
Db      110  CyProAspLeuValAspLysTyLeuGluThrSerIleArgTyValThrProArg 129
Qy      505  GGTAGCTGCGCTCGCGCGGATCAAGACAGTTCGAATTCGACAGAGGTCAAGTGG 564
Db      130  AlaLeuSerAspProGluProArgProGluProSerIleThrValProProHisTy 149
Qy      565  GAGACAGGCTCCCCCAACCAAG-----CGGAGGGCGGAGATCGGGGC 609
Db      150  ProValValCysCysProThrSerGlySerGlnMetCysSerLysArgProAlaArgGly 169
Qy      610  CGGCGCAGTGTGTGTCAGACCGCGCGGAGAGAGGGGCCACCGCCACCCAGCAGCAG 669
Db      170  ArgThrThrLeuGlySerGlnArgLysArgArg----- 180
Qy      670  TCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAACCTGTGACATCCGGCTCGGGTT 729
Db      181  ---LysSerValThrProAspProLysGluLysGlnThrCysAspIleArgLeuArgVal 199
Qy      730  CGAGCAGAGTACTGCGCAGATGGGCCAGCTTGAGACAGGCGCTGGCAGTCCCGCGGCC 789
Db      200  ArgAlaGluTyrcysGlnHisGluThrAlaLeuGlnGlyAsnValPheSerAsnLysGln 219
Qy      790  CAGGCGCTGGCGCGGAGCTGACGTGTTGGGCGAGGCGACCGCAGTGTGCTGCAAG 849
Db      220  AspProLeuGlnArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
Qy      850  GACCTGGGCTGTGTGTGTGTGACATCAAGTTCAGAGCTCTCTATCTGACGCTTTC 909
Db      240  AspLeuGlySerIleIleCysAspIleLysPheSerGlnLeuThrLysAspAlaPhe 259
Qy      910  TGGGGGACTACTGATGAGTGGCGCGCTGCGGCGGCGCTGCTTCCGACTGAG 969
Db      260  TrpArgAspLysTrileAsnGlySerLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
Qy      970  GCCCTCGAGAGAGCTGTGGGCGGAGAGGCTGTGCGCTGCTGCTGATGATGAGGCT 1029
Db      280  SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuValAsnValAspGlu 299
Qy      1030  GACTATGAGCTTGGCGCGCGCGCTGTG 1059
Db      300  AspTyrgLnuLeuGlyArgGlnLysLeuLeu 309

RESULT 7
US-09-109-273-6
; Sequence 6, Application US/09109273
; Patent No. 6063760
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emdad S.
; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
; NUMBER OF INVENTIONS: OF MAKING THE SAME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/109,273
; FILING DATE:

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/859,167
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TUI-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-109-273-6

```

```

Alignment Scores:
Pred. No.: 2,66e-38 Length: 318
Score: 641.50 Matches: 137
Percent Similarity: 62.90% Conservative: 58
Best Local Similarity: 44.19% Mismatches: 94
Query Match: 18.15% Indels: 21
DB: 3 Gaps: 5

```

```

US-10-030-271-3 (1-1883) x US-09-109-273-6 (1-318)
Qy      157  TGGAGAGAGAGAGAGAGGCTGAGCTACTAGCGGATGCTGCTTCAACGTAATGTTGAG 216
Db      12  TrpProGlnGlnArgGlyGluGlnGluLysLeuTySerLeuHisIleArgMetPheAsp 31
Qy      217  GTGTGTGGCGGCGCAACTGACCGAGTGCAGCTGAGACTCTGCTGCTTCTG---CTGAT 273
Db      32  IleValGlyThrHisLeuThrHisArgAspValArgValLeuSerPheLeuHisValAsp 51
Qy      274  GAGGCTCTGGGCGCGCGGAGGCTTAGCCCGGCGCGGAGGCGCTTACAGCTCTGCTG 333
Db      52  ValIleAspAspHisGluArgGlyLeu-----IleArgAsnGlyArgAspPheLeuLeu 69
Qy      334  GAGCTGAGAGCGCGCGGCGGAGTGCAGCGCGGAGCAACTGCGGCTGCGGCAACTCTG 393
Db      70  AlaLeuGlnArgGlnGlnLysArgCysAspGlnSerAsnPheArgGlnAlaLeuGlnLeu 89
Qy      394  CGCGTGTGCGCGCGCGCACGACTGTGCGCACTTGGCGCGGAGCGCGCGCGCAGTG 453
Db      90  ArgIleIleThrArgHisAspLeuProTyValThrLeuLysHisArgAlaVal 109
Qy      454  TCTCCA-----GACGCTATAGCTATGGCACTCCAGCTTCAAGAGACAGAG 504
Db      110  CyProAspLeuValAspLysTyLeuGluThrSerIleArgTyValThrProArg 129
Qy      505  GGTAGCTGCGCTCGCGCGGATCAAGACAGTTCGAATTCGACAGAGGTCAAGTGG 564
Db      130  AlaLeuSerAspProGluProArgProGluProSerIleThrValProProHisTy 149
Qy      565  GAGACAGGCTCCCCCAACCAAG-----CGGAGGGCGGAGATCGGGGC 609
Db      150  ProValValCysCysProThrSerGlySerGlnMetCysSerLysArgProAlaArgGly 169
Qy      610  CGGCGCAGTGTGTGTCAGACCGCGCGGAGAGAGGGGCCACCGCCACCCAGCAGCAG 669
Db      170  ArgThrThrLeuGlySerGlnArgLysArgArg----- 180
Qy      670  TCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAACCTGTGACATCCGGCTCGGGTT 729
Db      181  ---LysSerValThrProAspProLysGluLysGlnThrCysAspIleArgLeuArgVal 199
Qy      730  CGAGCAGAGTACTGCGCAGATGGGCCAGCTTGAGACAGGCGCTGGCAGTCCCGCGGCC 789
Db      200  ArgAlaGluTyrcysGlnHisGluThrAlaLeuGlnGlyAsnValPheSerAsnLysGln 219

```


QY 790 CAGCGCTGGCGCGCAGCTGACGTGTTGGGAGGCCAGCGAGTGGCTCAAG 849
 DB 220 AspProLeuGlnArgGlnPheGlnIleValThrIleLeuLysSerArg 239
 QY 850 GACCTGGGCTGCTGGTTTGACATCAAGTTCAGAGCTCTCTATCTGGAGCCTTC 909
 DB 240 AspLeuLysSerIleIleCysAspIleLysPheSerGluLeuThrIleAspAlaPhe 259
 QY 910 TGGGGCAGTACCTAGTGGCGCGCTGCTGCGAGCCCTGGCGGCGTTCTGAGCTGAG 969
 DB 260 TrpArgAspTyrIleLeuGlnSerLeuLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
 QY 970 GCCCTGCGAAGAGCTGTGGCGCGGAGAGCTGTGCTGCTGATGATGAGAGCT 1029
 DB 280 SerLeuLysGlnAlaValGlnIleGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
 QY 1030 GACTATGAGCTGGCGCGCGCGCTGTG 1059
 DB 300 AspTyrGlnLeuGlnArgGlnLysLeuLeu 309

RESULT 8

US-09-276-993-6
 ; Sequence 6, Application US/09276993
 ; Patent No. 6207801

GENERAL INFORMATION:

APPLICANT: Alnemri, Emdad S.
 APPLICANT: Fernandez-Alnemri, Teresa

TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
 TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS

NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESSES:

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6207801 is
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia

STATE: PA
 COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: WINDOWS

SOFTWARE: WordPerfect
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/276, 993

FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/859,167

FILING DATE:

ATTORNEY/AGENT INFORMATION:
 NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229
 REFERENCE/DOCKET NUMBER: TJU-

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
 LENGTH: 318 amino acids

TYPE: amino acid
 TOPOLOGY: linear

MOLECULE TYPE: protein
 US-09-276-993-6

Alignment Scores:
 Pred. No.: 2,66e-38

Score: 641.50
 Percent Similarity: 62.90%

Best Local Similarity: 44.19%
 Query Match: 18.15%

DB: 3
 Gaps: 5

US-10-030-271-3 (1-1883) x US-09-276-993-6 (1-318)
 QY 157 TGGGAGGAGATGATGCTGCTGACCTAGTACGAGATGCTGCTTACCGATGCTGAG 216
 DB 12 TrpProGlnGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 31
 QY 217 GTGGTGGGCGGCGCACTGACCGAGTGGAGCTGAGCTCTGCGCTTCTG--CTGAT 273
 DB 32 IleValGlnIleThrIleLeuThrIleValArgAspValArgValLeuSerPheLeuPheValAsp 51
 QY 274 GAGGCTCTGCGCGCGCGGAGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 333
 DB 52 ValIleAspAspIleGlnArgGlnLeu-----IleArgAsnGlnGlnArgPheLeuLeu 69
 QY 334 GAGCTGAGGCGCGCGCGCGAGTGGCGGAGAGCACTGCGGCTGCGGCGCACTCTG 393
 DB 70 AlaLeuGlnArgGlnGlnArgCysAspGlnSerAsnAsnAsnAsnAsnAsnAsnAsnAsn 89
 QY 394 CGCGCTGCG 453
 DB 90 ArgIleIleThrIleThrIleAspLeuLeuProTyrValThrLeuLysValArgAlaVal 109
 QY 454 TCTCCA-----GAGCGCTATGCTATGCGACCTTCCAGCTTTCAAGAGACGACAG 504
 DB 110 CysProAspLeuValAspLysTyrLeuGlnGlnIleThrSerIleArgTyrValThrProArg 129
 QY 505 GGTAGTGGCG 564
 DB 130 AlaLeuSerAspProGlnProArgProGlnProGlnProSerLysThrValProProIleTyr 149
 QY 565 GAGACAGGCTCCCG 609
 DB 150 ProValValCysCysProThrSerGlnSerGlnMetCysSerLysValArgProAlaArgGly 169
 QY 610 CGGCGCGAGTGTGTGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 669
 DB 170 ArgThrThrLeuGlnSerGlnArgLysArgArg----- 180
 QY 670 TCAGAGCG 729
 DB 181 ---LysSerValThrProAspProLysGlnLysGlnThrCysAspIleArgLeuArgVal 199
 QY 730 CGAGCAGTACTGCGAGCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 789
 DB 200 ArgAlaGlnIleThrCysGlnIleGlnIleThrAlaLeuGlnIleValAsnValPheSerAsnLysGln 219
 QY 790 CAGCGCTGCG 849
 DB 220 AspProLeuGlnArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
 QY 850 GACCTGGGCTGCTGGTTTGACATCAAGTTCAGAGCTCTCTATCTGGAGCCTTC 909
 DB 240 AspLeuLysSerIleIleCysAspIleLysPheSerGluLeuThrIleAspAlaPhe 259
 QY 910 TGGGGCAGTACCTAGTGGCGCGCTGCTGCGAGCCCTGGCGGCGTTCTGAGCTGAG 969
 DB 260 TrpArgAspTyrIleLeuGlnSerLeuLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
 QY 970 GCCCTGCGAAGAGCTGTGGCGCGGAGAGCTGTGCTGCTGATGATGAGAGCT 1029
 DB 280 SerLeuLysGlnAlaValGlnIleGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
 QY 1030 GACTATGAGCTGGCGCGCGCGCTGTG 1059
 DB 300 AspTyrGlnLeuGlnArgGlnLysLeuLeu 309

RESULT 9

US-09-723-450-6
 ; Sequence 6, Application US/09723450
 ; Patent No. 6576751

GENERAL INFORMATION:
 APPLICANT: Alnemri, Emdad S.

TITLE OF INVENTION: FADD-Like Anti-Apoptotic Molecules, Methods Of Using The Same, And

QY 454 TCTCCA-----GAACGCTATAGCTATGACCTTCCAGCTTCTTCAAGAGACAGAG 504
Db 110 CyseProAspLeuValAspLysTyrLeuGlnGlnThrsSerIleArgTyrValThrsProArg 129
QY 505 GGTAGCTCCGCTCGCCGCTCGGACAGTTCGAAATTCCTCAGCAGGAGTCACTGG 564
Db 130 AlaLeuSerAspProGlnProArgProGlnProSerIleArgTyrValProProIleTyr 149
QY 565 GAGACAGAGCTCCCGCCAGCCAG-----CGGACAGCGGAGAGTCCGAGG 609
Db 150 ProValValCysCysProThrsSerGlySerGlnMetCysSerIleArgProIleArgGly 169
QY 610 CGGCGCAGTGTGTGTCAGAGCGGCGGAGAGGAGGCGCCAGCGCAGCCAGCAGAG 669
Db 170 ArgThrThrLeuGlnSerGlnArgLysArgLys-----LysSerValThrsPro 180
QY 670 TCAGAGCCCGCCAGACCTTCTCTGAAAGGCAAGTGAACCTGTGACATCCGGCTCCGGTT 729
Db 181 ---LysSerValThrsProAspProLysGlnLysGlnThrsCysAspIleArgLeuArgVal 199
QY 730 CGAGCAGAGTACTGCGAGCATGGGCGACGCTTGGAGCAGGCGCGGAGATCCCGCGGCCC 789
Db 200 ArgAlaLeuTyrCysGlnHisGlnThrsAlaLeuGlnGlnValPheSerAsnLysGln 219
QY 790 CAGGCGCTGCGCGCAGCTGAGCTGTGGGCGAGCGCACCGCAGTGTGCTGCTCAAGG 849
Db 220 AspProLeuGlnArgLysPheGlnPheGlnValPheAsnThrIleLeuLysSerArg 239
QY 850 GACCTGGGCTGTGTGTTGTGATCATCAAGTTCACAGCTCTCTTATTCGAGCGCTTC 909
Db 240 AspLeuLysSerIleIleCysAspIleLysPheSerGlnLeuThrTyrLeuAspAlaPhe 259
QY 910 TGGGGGACACTACCTGAGTGGGCGGCTGTGAGCGGCGGCGGCTTCTGACAGG 969
Db 260 TTrpArgAspTyrIleAsnGlnSerLeuLeuGlnValLeuLysGlnValPheIleThrAsp 279
QY 970 GCCCTGCGAGAGGCTGTGGGCGGAGAGCTGTGCTGCTGATGATGATGAGGCT 1029
Db 280 SerLeuLysGlnAlaValGlnHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
QY 1030 GACTATGAGGCTGCGCGGCGGCGGCTGTG 1059
Db 300 AspTyrGlnLeuGlnArgGlnLysLeuLeu 309

RESULT 11
US-09-733-167A-6
Sequence 6, Application US/09733167A
Patent No. 6696547
GENERAL INFORMATION:
APPLICANT: Peter, Marcus
APPLICANT: Krammer, Peter
TITLE OF INVENTION: Protein for Regulation of Apoptosis
FILE REFERENCE: 4121-120
CURRENT APPLICATION NUMBER: US/09/733.167A
CURRENT FILING DATE: 2000-12-08
PCT/DE99/01712
PRIOR APPLICATION NUMBER: PCT/DE99/01712
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
PRIOR FILING DATE: 1998-06-08
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 210
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: C-DEBD, which is a deletion mutant of human DEBD comprising amin
OTHER INFORMATION: acids 109-318 of the naturally occurring huma DEBD.
US-09-733-167A-6

Alignment Scores: 6.59e-25 Length: 210
Pred. No.: 455.00 Matches: 87

Percent Similarity: 72.56% Conservative: 32
Best Local Similarity: 53.05% Mismatches: 35
Query Match: 12.87% Indels: 10
DB: 4 Gaps: 1

US-10-030-271-3 (1-1883) x US-09-733-167A-6 (1-210)

QY 568 ACAGAGCTCCCGCCCAACAGCGGAGCGGAGTCCGAGGCGGCGCCAGTGTGTGCC 627
Db 48 ThrSerGlyProGlnMetCysSerIleArgProAlaArgIleArgIleThrLeuGlySer 67
QY 628 AGACGCGCGGAGAGAGGAGGCGCCAGCCGACCCAGCAGCAGTCAAGACCCGCGACCT 687
Db 68 GlnArgLysArgArg-----LysSerValThrsPro 77
QY 688 TCTCTGAGGCAAGAGTGAACCTGTGACATCCGGCTCCGGTTGAGCAGAGTCACTGGAG 747
Db 78 AspProLysGlnLysGlnThrCysAspIleArgLeuArgValArgAlaGlnTyrCysGln 97
QY 748 CATGGGCGACGCTTGGAGCAGGCGCGTGGACATCCCGGCGGCGCCAGCGGCGGCGAG 807
Db 98 HisGlnThrAlaLeuGlnGlnValPheSerAsnLysGlnAspProLeuGlnArgGln 117
QY 808 CTGAGCGTGTGGGCGAGCGCACCGCAGTGTGCTGCTCAAGAGGACCTGTGCTGTG 867
Db 118 PheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArgAspLeuGlySerIle 137
QY 868 TGTGACATCAAGTTCACAGCTCTCTATCTGAGCGCTTTCGGGCGGCTACCTGAGT 927
Db 138 CysAspIleLysPheSerGlnLeuThrTyrLeuAspAlaPheThrArgAspTyrIleAsn 157
QY 928 GGGCGCTGTGAGGCGGCTGTGAGGCGGCTGTGCTGCTGAGGCGGCTGTGAGGAGCTGTG 987
Db 158 GlySerLeuLeuGlnAlaLeuLysGlnValPheIleThrAspSerLeuLysGlnAlaVal 177
QY 988 GGGCGGAGGCTGTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1047
Db 178 GlnHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlnLysTyrGlnLeuGlnArg 197
QY 1048 CGCGGCTGTG 1059
Db 198 GlnLysLeuLeu 201

RESULT 12
US-09-252-991A-28408
Sequence 28408, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
APPLICANT: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074.788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094.190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 28408
LENGTH: 1411
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28408

Alignment Scores: 1.56e-12 Length: 1411
Pred. No.: 286.00 Matches: 173
Percent Similarity: 32.70% Conservative: 35
Best Local Similarity: 27.20% Mismatches: 202
Query Match: 8.34% Indels: 226
DB: 4 Gaps: 32

US-10-030-271-3 (1-1883) X US-09-252-991A-28408 (1-1411)

[illegible]

```

Oy      583  -----TTGGGGGGGAGAGCCTGTCCCTCCAGTCAACCTGGTGAAGAAAT 545
Db      300  ArgGAGLeuProIleuLeuYsValIProIyGlyuSenSerProIaAlaGlnThIinSer 319
Oy      544  TTGCAGAACTGCTTGACT-----GCCAGCGCGAGCGAGCTAC 506
Db      320  AsnGInGlySerProThiMerPheArMetSerAspAlaAlaGluAspGlyLeuThi 339
Oy      505  -----CCTGTGCTCTCTTTGAAGAGC---TGGAGTGCCTAGCTGTAAAGGTT 461
Db      340  LeuAlaGlnAspProAlaThiGlyLeuYsAlaIleIleAlaIleThiSerThiArgLeu 359
Oy      460  CTGGAAGACACTGGCGCGCGCGCTTGGCGCGCA-----GGTGGCGAGACAGT 413
Db      360  GLyProAlaLeuGlyGlyCyAsArgTyGlyLeuProTyGProAsnAspGlnAlaIleGly 379
Oy      412  -----CGTGGCGGGCGAGACACGCGAGAGTT 386
Db      380  AspAlaHisPProVroGlyProGlyuHisGlyLeuGlnGlyCyProGlyArgAlaGlyAla 399
Oy      385  GCC-----CCAGAGCGCGAGGTGCTGTCCGCGCGACATGCC 350
Db      400  GLyArgArgGlnGlyCyLAsnHisPProVroAlaAlaPProGlyGlnSerAspArgAlaVal 419
Oy      349  CGCGGGGCTCCAGCT-----CCAGAGAGAGCTTAAAGCGCGCTGGCGCGCGCG 302
Db      420  ArgGlyLeuAlaGlnAlaLeuHisAsArgTySerProAlaArgAlaGluHisHisCyAsGlyGln 439
Oy      301  CTAAAGCTCCGGCGG-----CGCGAGAGACT 275
Db      440  ArgTyGlyLeuGlnArgHisGlyGlyLeuHisAsArgProAlaAspProVroArgHisGlnHis 459
Oy      274  CATCCAGCA----- 266
Db      460  HisPProGlyArgArgProLeuAlaAlaIArgArgProAlaGluValArgGluAspSerArg 479
Oy      265  GAAAGCGCAGAGACTCCAGCTGCAGCTCGG-----TCAGTT 230
Db      480  GlnArgProGlyAlaArgProGlyGlnArgArgProGlyArgProAlaGlyArgGlyAlaGly 499
Oy      229  GCGCGCGCCACACTGGAACATACGCTGAAGGAGACAGATCC-----CGTAGTAGTCCA 176
Db      500  AlaArgProArgArgGlyLeuArgGlnIleGlyArgAlaThiGlyGlyArgAspArgAlaAla 519
Oy      175  GGCACCTATCTCTCTCCAGCAGCGAGCGCGGGGTGCAGCCGATATAGCGCATTTCCCGGG 116
Db      520  GLyLeuArgProArgSerArgProGlyAlaThiGlyArgAlaGlnAlaArgCyProSer 539
Oy      115  GAGGGA---GCGGAAAGACTCAGAACCCGCGCGAGTTGGCGTCCGAGTCTCTGG 59
Db      540  AlAGlyAlAGlyGlyLeu-AlaLeuHisPProLeuArgHisPProArgAlaLeuArgProGly 559
Oy      58  GCGAGCGCGG-----CAGCGACTGTCTTTG 34
Db      559  yArgArgAlaHisGlnPProGlyArgGlnProValaLeu 572

RESULT 13
US-09-219-849-3
; Sequence 3, Application US/09219849
; Patent No. 6150081
; GENERAL INFORMATION:
; APPLICANT: VAN HEERDE, GEORGE V.
; APPLICANT: VAN RIJN, ALEXIS C.
; APPLICANT: BOUTWISTRA, JAN B.
; APPLICANT: DE WOLF, FREDERIK A.
; APPLICANT: MOORBROEK, ANDREAS
; APPLICANT: WERTEN, MARC W.T.
; APPLICANT: WIND, RICHÈLE D.
; APPLICANT: VAN DEN BOSCH, TANJA J.
; TITLE OF INVENTION: SILVER HALIDE EMULSIONS WITH RECOMBINANT COLLAGEN
; TITLE OF INVENTION: SUITABLE FOR PHOTOGRAPHIC APPLICATION AND ALSO THE
; TITLE OF INVENTION: PREPARATION THEREOF

```


; INFORMATION FOR SEQ ID NO: 52:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 561 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-642-255-52

Alignment Scores:

Pred. No.:	2,28e-11	Length:	561
Score:	268.50	Matches:	139
Percent Similarity:	34.41%	Conservative:	10
Best Local Similarity:	32.10%	Mismatches:	153
Query Match:	7.83%	Indels:	132
DB:	1	Gaps:	22

US-10-030-271-3 (1-1883) x US-08-642-255-52 (1-561)

```

QY 1302 GGGCCAGCC-----CCGCGCTCCGAGGCTAAGGGGCGACACTTGGGGGTAGAGAGCT 1249
    |||||
DB 190 G|PProAlaG|PProProG|yAlaPProG|yProAlaG|yPProProG|yAlaPProG|yPro 209
    |||||
QY 1248 CTGGGAGGAAACTAGCTGGAATGTTAGGCTCAGAGCCGACATCTGTGGGA---- 1192
    |||||
DB 210 Alagly-----ProProG|yAlaPProG|yProAlaG|yPro 221
    |||||
QY 1191 ---GGGCGTCAAGGGGTAGAGATGTGCTCTCCAGAGAAAGGTGGCCGAGACTT 1135
    |||||
DB 222 ProG|yAlaPPro----- 225
    |||||
QY 1134 GGAGGTGGATCATCTGCGAGTCTGTGATCAGAGGCTCTGTGCGGCGCGCC-CCC 1076
    |||||
DB 226 G|PProAlaG|PProProG|yAlaPProG|yProAlaG|yPProProG|yAlaPProG|yPro 245
    |||||
QY 1075 CTTCCTCTCCATCAGCAACAGCGCGCGCGGCGAGCTCATAGT----- 1031
    |||||
DB 246 AlaglyPProProG|yAlaPProG|yProAlaG|yPProProG|yAlaPProG|yProAlaG|y 265
    |||||
QY 1030 CAGCCTCATTCACACTGACGACGAGCGAGCAAGCCTCCGCGCCACAGCCTCTCGCAGGG 971
    |||||
DB 266 ProProG|yAlaPProG|yProAlaG|y-----ProLyG|yAlaPProG|yProAlaG|y 283
    |||||
QY 970 CCTCACTCAGAAACACGCCCGCGAGGCTGTGACGAGGGCGCCACTCAGTAGTCCGCC 911
    |||||
DB 284 ProLyG|yAlaPProG|yProAlaG|yPro--LyG|yAlaPProG|yProAlaG|yPro 302
    |||||
QY 910 AGAAGGCGTCCGATAGAGAGCTGTGAGAACTTATGTCACAAACACAGAGCCAGGT 851
    |||||
DB 303 LyG|yAlaPProG|yProAlaG|y-----ProProG|yAlaPProG|y 316
    |||||
QY 850 CCCTTGAGGCGACACTGCGGTGCGCTGCGCAACAGTTCAGCTGCCGCCA--GCG 794
    |||||
DB 317 ProAlaG|yPro-----ProG|yAlaPProG|yProAlaG|yProProG|yAla 332
    |||||
QY 793 CCTGGGCGCGCGGATGCCA--CGCCTGTCTCAAGCTGCGCCATCGTCGCACTACT 737
    |||||
DB 333 ProG|yAlaPProG|yProProG|yAlaPProG|yPro--AlaG|yPro----- 347
    |||||
QY 736 CTGCTCGAACCAGGACCGGATGTCACAGTCACTTGGCTTCAGAGAAAGTCTGCGG 677
    |||||
DB 348 -----ProG|yAlaPProG|yProAlaG 355
    |||||
QY 676 GCTCTGACTGCTGCGGTGCGGCTGCGGCGCTCTCGCGCGCGCTGCGACCA--C 620
    |||||
DB 355 LyPProProG|yAlaPProG|yProAlaG|yPProPro-----G|yAlaPProG|yP 371
    |||||
QY 619 CACTGGGCGCGCGCGCACTCCCGCGCTGCTGTTGGGGGAGGCTGTCTCCACT 560
    |||||
DB 371 roAlaG|yPPro-----G|yAlaPPro-G|yProAla 381
    |||||
QY 559 GACCTGCTGAGAAATTGGCAAACTGTGACTGCCGAGGCGACGCGACGTAACCTCTG 500
    |||||
  
```

```

DB 382 G|yPProPro-----G|yAlaPProG|yPro----- 389
QY 499 TCCTCTTGAAGAGCTGAGAGTGCCATAGCTATTAGCGTTTGAGACACTGCGCGCGCC 440
    |||||
DB 390 -----AlaG|yPro 392
    |||||
QY 439 GCTTGGCGCGCGAGTGTGCTGCGCGCACTGCGCGCGCGCGCGCGAGT----- 386
    |||||
DB 393 ProG|yAlaPProG|yProAlaG|yPProProG|yAlaPProG|yProAlaG|yPProProG|y 412
    |||||
QY 385 GCGCCAGAGCGCGAGTGTGCTGCGCGCACTGCGCGCGCGCGCGCGCTCAGCTCAGCAGGA 326
    |||||
DB 413 AlaPProG|yProAlaG|y-----ProProG|yAlaPProG|yProAlaG|y 427
    |||||
QY 325 GCTTAGCGCGGTGCGCGCGCGCGCGTAAAGCTCCGCGCGCGCGCGAGACT----- 275
    |||||
DB 428 ProProG|y--AlaPProG|yPro--AlaG|yPProProG|yAlaPProG|y--ProAlaG|yPro 446
    |||||
QY 274 -----CATCCAGCAAAAGCCAGAGCTCCAGCTCGCACTCGTCACTTGGC 227
    |||||
DB 447 ProG|yAlaPProG|yProAlaG|yPProProG|yAlaPProG|yProAlaG|yPProProG|y 466
    |||||
QY 226 CGCCACCACTCGAACAATAGCGTGAAGGACAGCATCCGTAATGTCAGGCACTCAT 167
    |||||
DB 467 AlaPProG|yProAlaG|yPProProG|yAlaPProG|yProAlaG|yPProProG|yAla--- 485
    |||||
QY 166 CCTCTCCAGCAGCGGGCGGGGTGACCCGATAGCGCCATTCGCGGGGAGGAGGC 107
    |||||
DB 486 ProG|yProAlaG|yPProProG|yAlaPProG|yProAlaG|yPProProG|yAlaPProG|y 505
    |||||
QY 106 -----GGAACAAGCTCAGAACCCGCGC 86
    |||||
DB 506 ProAlaG|yPProProG|yAlaPProG|y 514
    |||||
RESULT 15
US-09-252-991A-20675
; Sequence 20675, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20675
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20675
Alignment Scores:
Pred. No.: 6.55e-11 Length: 726
Score: 262.50 Matches: 185
Percent Similarity: 30.34% Conservative: 31
Best Local Similarity: 25.98% Mismatches: 232
Query Match: 7.65% Indels: 264
DB: 4 Gaps: 33
US-10-030-271-3 (1-1883) x US-09-252-991A-20675 (1-726)
QY 1541 GCTGAAGCCAGCTGTGTGCTCA-----ATTAGAGCT 1506
    |||||
DB 51 AlagluG|nPProG|yG|yArgG|yG|nAlaArgTnArGProAlaPProThrThxArgG|y 70
    |||||
QY 1505 CCTGAGAGAGAGCCCTTAGCGCAAAAGATGTGCTGCGAAGCAGCGCAGGAGAGG 1446
    |||||
DB 71 ProHisPProG|yLeuArgArgArgValLeuAlaTTPG|nAspArgAlaAspG|nG|n 90
    |||||
  
```

```

QY 1445 CCC-----AGGAGAATCTCAGCAGGGCAGTGTGACATGAGGAGTGTGTGCAC 1395
Db 91 ProValLeuArgGluLeuArgProAlaHisAlaValProArgGluAlaHisHis 110
QY 1394 CTGACAGATGACACAGCCCAAGATCAAGATGATGTGACAGGGGCTTTGGTAC 1335
Db 111 ValProHisArgAlaValProArg-----SerAlaArgGluAlaGlnLeuHis 126
QY 1334 TGTGTAGTGTGCTCTCTCTGAGATACAGGCCAGCCCGCTCCGAGAGCTAAG 1275
Db 127 ProProAlaAlaAraArgAraArgAraProGlyArgArgGlnArgProValGlnAlaArg 146
QY 1274 GGGGACACCTGGGAGTGAAGAGAGTGGAGAGAACTCAGCTGAAATGTTTAGGC 1215
Db 147 AlaAlaProLeuGlyGlnHisSerAlaGlyHisArgArgProGlyGlnHisGlyProGly 166
QY 1214 CTCGAGCCCACTCTCTGTGGAGGGCTGTCAAGGGGTAGAGATGCTCTCCACAG 1155
Db 167 LeuArg-ProGlyArgGlnArgGlnGlyAlaAlaGly-----Gly 179
QY 1154 AGAAGTGGCCCGAGACTTGAGATGGAGATCAATCCGCGCACTCTGATCAGAGAGCC 1095
Db 179 YThrGlyAraProAla-----ArgLeuProProGlyCyserProGlySerArgArgAla 197
QY 1094 TCTGTCCGCGCCGCGCCCT-----CTTCTCTCTCCATCAGC 1059
Db 197 ArgArgProGlyAlaArgProGlyLeuAlaProArgProGlyGlnLeuArgProProAla 217
QY 1058 AACGCGGCGCGCCGCGCCCT---CATAGTCAGCTCAT---CACAATGACACAGAG 1005
Db 217 AlaArgProGlyProAlaAraProArgHisAlaArgProGlyCyserProGlnArgAlaGly 237
QY 1004 CGAAGAGCT-----CCGCGC---CCACAGCT----- 980
Db 237 ValAraProValAraAlaAlaGlnArgProGlyAraProValProAlaValArgArg 257
QY 979 -----CTGCAAGGCTCTCAGTCAAGAAC 956
Db 257 GATGCTHisArgGluGlnHisGlyGlyLeuAlaGlnValHisProAlaThrAlaArgGly 277
QY 955 -CGCGCGCGAGGCTCTGACAGAGCGCCCACTCAGTATGCTCCGCGCGAGAGCTCCAGA 897
Db 277 ArgArgProGlnArgProValArgArgAlaGlnArgAraArgAlaLeuGlyArgProHis 297
QY 896 TAGGAGAGCTCTGAGACTTGATGTCAAAACA----- 863
Db 297 ArgArgAlaLeuArgGlnArgHisArgProAraProGlyArgHisGlyProProAl 317
QY 863 ----- 863
Db 317 ArgArgProAlaAlaArgArgArgAlaAlaValAlaGlyGlnAlaAlaGlnGlnGly 337
QY 862 -----CAGAGCCCAAGCTCTCTG---AGCGCAGACT 834
Db 337 ArgArgArgGlnGlyAlaAlaArgGlnHisArgProGlySerAlaGlySerAlaAlaLe 357
QY 833 GCGG----- 830
Db 357 ValArgAlaHisArgArgAlaGlnGlySerAlaAlaGlyGlyLeuArgAlaArgProGlyPr 377
QY 829 -TGGCTCTGCGCAACAGCTCAGCT---GCGCGCGCAGAGCTGGGCGCGC----- 782
Db 377 ArgProAlaAlaGlnGlnProAlaSerAlaAlaArgProGlyGlnGlnArgGlyTh 397
QY 782 ----- 782
Db 397 ValArgAlaAlaArgArgGlnAlaAlaArgGlnArgProThrAlaValArgThrAra 417
QY 781 -----GGGATGCGCAGC-----CC 768
Db 417 GGUArgGlnGlyArgProArgProProProProProAlaGlnProGlnAraGlnPr 437

```

```

QY 767 TGTCTCAAGGCTGGCCCATCTGCGAGTACTCTGTCTGAACCCGAGCCGAGTCTCAG 708
Db 437 AlaProAlaAlaAlaArgProGlyArgAra--AlaAlaArgAraAraProGlnArgThrHis 456
QY 707 GTCATTTGCTTCAAGGAAAGCTTGGCGGCTCTACAGTCTGCTGGGCTGGCTGGG 648
Db 457 GlyThrAlaProAlaGln-----GlyAlaAla 465
QY 647 GCGCTCTCGCGCC-----CGTCTGACACA-----CCA 618
Db 466 GlnProGlnArgGlnProAraArgHisGlyArgHisGlnProGlyAraProGlyValLeuPro 485
QY 617 CTGGGC-----CGCGCCGACCTCCGCTCCGCTGCTGTGGAGGAGAGCTGTCTCC 564
Db 486 ArgGlyGlyAlaArgProArgGlnProArgAlaArgGlyArgHisGlyGlnPro----- 503
QY 563 CACTGACCTCTGCGAATTTGCAAGACTCTGACTGCGGACGCGACGCGACCTACCC 504
Db 504 -----HisGlyArgArgHisAraPro 510
QY 503 TCTGTCTCTTTGAAGAGTGAAGTGCATAGACTATAGCGTTCTGAGACACTGGCGCG 444
Db 511 -----ProAlaGlnArgGlyGly 516
QY 443 CGCGCTTGGCGCGCA-----GGTGGCGACAGAGTCTGGCGCGCGACAGCGCGCA-- 392
Db 517 ProAlaLeuArg-ProArgArgAlaAlaAlaThrGlnProAlaLeuArgProAlaArgThrHis 536
QY 391 -----GGAGTTCGCCAGACAGCGCGAGTTGCTCTCGCGC-- 356
Db 536 ValArgGlnGlnGlnGlyArgProValProProAlaAlaGlnAraArgAlaArgGly 556
QY 355 -----ACTGCGCGCGCGCTCCAGCTCC 333
Db 556 ArgProAlaAraProAlaLeuLeuArgHisArgProGlyAlaGlyGlyArgProAlaArg 576
QY 332 AGGAGAGCTTAGGCGCGCTGGCGCGCGCTAAGCTCCGCGCGCGCGACAGAGCTCA 273
Db 576 AlaProAlaArgGlyArgHisArgAlaLeuGly-----GlnArgArgAlaAraAla 593
QY 272 TCCAGCGAAGAGCGAGAGCTCAGCTCGCATCTGCTGAGTTGCTCCGCGCGCA 218
Db 593 AaArgAlaAlaHisArgProGlyArgGlnThrValAlaArgGlyAraProHisAlaArgProGlnGly 613
QY 217 -----CCTCGAACAATAGGAGTGAAGCGACAGACTCCCG 186
Db 613 ArgProGlyArgProAlaAlaGlyThrProAlaGlnProGlyLeuArgArgValArgArg 633
QY 185 TAGTAGTCAAGGACT-----CATCTCTCTCCAGCAGCGGCGCGG 144
Db 633 GGUAlaAlaGlyThrValAlaArgSerArgArgHisArgProProProAlaAlaProPro-- 652
QY 143 GTGACCCGGAATAGCGCA---TTCCCGGAGAGAGAGCGGAAACAAGCTCAAGAACCGG 87
Db 653 -----LeuAlaProAlaAlaArgProGlyGlnGlyGly-----GlnProAl 665
QY 86 CAGATTCTGCGTCGAGATCACTCTGGGAGAGCGCGCAGCAGCTGTTTGATCTTT 27
Db 665 ArgGlyAlaAlaArg-LeuThrProLeuProArgArgAlaProAlaArgLeuProLeuG 685
QY 26 CTGGCAGCTTCTCTATATTAGCCT 1
Db 685 ValAlaThrLeuAraArgLeuAlaPro 693

```

Search completed: March 22, 2005, 01:57:45
 Job time : 69.9542 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 23, 2005, 09:08:14 ; Search time 164.414 Seconds
(without alignments)
9046.504 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctaccgcggctcgac.....tcagtcgtgatgagctgac 909

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues 2405568

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents NA:*
2: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfill1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	541	59.5	1067	4 US-09-620-312D-853	Sequence 853, App
2	193.8	21.3	1142	4 US-09-733-167A-2	Sequence 2, Appl
3	192.2	21.1	1045	3 US-08-859-167-3	Sequence 3, Appl
4	192.2	21.1	1045	3 US-09-109-273-3	Sequence 3, Appl
5	192.2	21.1	1045	3 US-09-276-993-3	Sequence 3, Appl
6	192.2	21.1	1045	4 US-09-723-450-3	Sequence 3, Appl
7	190.6	21.0	1142	4 US-09-733-167A-4	Sequence 4, Appl
8	190.6	21.0	1200	3 US-08-859-167-5	Sequence 5, Appl
9	190.6	21.0	1200	3 US-09-109-273-5	Sequence 5, Appl
10	190.6	21.0	1200	3 US-09-276-993-5	Sequence 5, Appl
11	190.6	21.0	1200	4 US-09-723-450-5	Sequence 5, Appl
12	146.2	16.1	630	4 US-09-733-167A-8	Sequence 8, Appl
13	88.4	9.7	342	4 US-09-733-167A-7	Sequence 7, Appl
14	50.8	5.6	1151	1 US-07-704-288C-2	Sequence 2, Appl
15	50.8	5.6	1151	1 US-08-093-372-1	Sequence 1, Appl
16	50.8	5.6	1151	1 US-08-379-259-2	Sequence 2, Appl
17	49.8	5.5	999	4 US-09-902-540-3684	Sequence 3684, Ap
18	49.8	5.5	16924	4 US-09-902-540-1178	Sequence 1178, Ap
19	49.8	5.4	929	4 US-09-902-540-3093	Sequence 3093, Ap
20	48.2	5.3	15789	4 US-09-902-540-1139	Sequence 1139, Ap
21	48.2	5.3	861	4 US-09-902-540-7313	Sequence 7313, Ap
22	48.2	5.3	5764	4 US-09-902-540-693	Sequence 693, Ap
23	47	5.2	603	4 US-09-902-540-9070	Sequence 9070, Ap
24	47	5.2	7719	4 US-09-902-540-969	Sequence 969, Ap
25	46.8	5.1	774	4 US-09-266-965-43	Sequence 43, Appl
26	46.8	5.1	1515	4 US-09-902-540-8120	Sequence 8120, Ap
27	46.8	5.1	7846	4 US-09-902-540-830	Sequence 830, App

C	28	46.8	5.1	53500	4 US-09-266-965-76	Sequence 76, Appl
	29	46.6	5.1	1509	4 US-09-724-797-89	Sequence 89, Appl
	30	46.4	5.1	9937	4 US-09-902-540-946	Sequence 946, Appl
	31	46.2	5.1	663	4 US-09-902-540-7269	Sequence 7269, Ap
	32	46.2	5.1	4324	4 US-09-902-540-684	Sequence 684, Appl
	33	45.8	5.0	699	4 US-09-902-540-9109	Sequence 9109, Ap
	34	45.8	5.0	831	4 US-09-252-991A-1479	Sequence 1479, Ap
	35	45.8	5.0	1962	4 US-09-252-991A-1693	Sequence 1693, Ap
	36	45.8	5.0	11963	4 US-08-804-227C-1	Sequence 1072, Ap
	37	45.8	5.0	43280	2 US-08-804-227C-1	Sequence 1, Appl
	38	45.6	5.0	601	4 US-09-949-016-19457	Sequence 19457, A
	39	45.6	5.0	601	4 US-09-949-016-19457	Sequence 19457, A
	40	45.6	5.0	601	4 US-09-949-016-49701	Sequence 49701, A
	41	45.6	5.0	601	4 US-09-949-016-49702	Sequence 49702, A
	42	45.6	5.0	126200	4 US-09-949-016-11824	Sequence 11824, A
	43	45.6	5.0	126200	4 US-09-949-016-13193	Sequence 13193, A
	44	45.2	5.0	2635	3 US-09-126-280-3	Sequence 3, Appl
	45	45.2	5.0	2670	3 US-09-126-280-1	Sequence 1, Appl

ALIGNMENTS

```

RESULT 1
US-09-620-312D-853
; Sequence 853, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundang
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhilwei
; APPLICANT: John Tillghast
; APPLICANT: Drmanac, Radote T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_Fl_genes Version 1.0
; SEQ ID NO 853
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (151)...(984)
US-09-620-312D-853

Query Match 59.5%; Score 541; DB 4; Length 1067;
Best Local Similarity 95.8%; Pred. No. 8.7e-113;
Matches 572; Conservative 0; Mismatches 10; Indels 15; Gaps 1;
QY 1 ATGGCGCTATCCGGGTCGACCCCGGCCGCTGTGGAGAGAGATAGTCCCTGACTAC 60
DB 151 ATGGCGCTATCCGGGTCGACCCCGGCCGCTGTGGAGAGAGATAGTCCCTGACTAC 210

```

QY 61 TACGGGATGCTGCTGCTTCAACCGTATGTTTCAGAGTGCTGAGGCGGCACTGACCGAGTGC 120
DB 211 TACGGGATGCTGCTGCTTCAACCGTATGTTTCAGAGTGCTGAGGCGGCACTGACCGAGTGC 270
QY 121 GAGCTGAGAGCTCTGAGCTTTTGTGCTGATGAGGCTCTGAGGCGGCGGAGGCTTAGCC 180
DB 271 GAGCTGAGAGCTCTGAGCTTTTGTGCTGATGAGGCTCTGAGGCGGCGGAGGCTTAGCC 330
QY 181 CGGGCCCGCAGCGGCTTGAAGCTCTGCTGAGCTGAGAGCGGCGGCGGAGTGCAGG 240
DB 331 CGGGCCCGCAGCGGCTTGAAGCTCTGCTGAGCTGAGAGCGGCGGCGGAGTGCAGG 390
QY 241 AGCAACCTGCGGCTGCTGAGGCACTCTGCGCGGCTGCTGAGGCGGCGGAGTGCAGG 300
DB 391 AGCAACCTGCGGCTGCTGAGGCACTCTGCGCGGCTGCTGAGGCGGCGGAGTGCAGG 450
QY 301 CACCTGCGGCGGCAAGCGGCGGCGGCGGCAAGTCTCCAGAAAGCTATAGCTAGGCACTCC 360
DB 451 CACCTGCGGCGGCAAGCGGCGGCGGCGGCAAGTCTCCAGAAAGCTATAGCTAGGCACTCC 510
QY 351 AGCTTTCAAGAGAGCAAGAGGAGTGCCTGCTGCGCGGCTGCGGAGTCAAGCACTTTCGA 420
DB 511 AGCTTTCAAGAGAGCAAGAGGAGTGCCTGCTGCGCGGCTGCGGAGTCAAGCACTTTCGA 570
QY 421 AATTCTCAGCAGGAGTCAAGGAGAGAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
DB 571 AATTCTCAGCAG-----GCTCTCCCCCAACCAAGCGGAGCGGAGT 615
QY 481 CGGGCCCGGCGGAGTGTGTGCAAGAGCGGCGGAGAGGCGGCGGAGCGGAGT 540
DB 616 CGGGCCCGGCGGAGTGTGTGCAAGAGCGGCGGAGAGGCGGCGGAGCGGAGT 675
QY 541 CAGCAGTCAAGAGCGGCGGAGAGCTTCTCTGAAAGCAAGTGAAGTCAAGTCCG 597
DB 676 CAGCAGTCAAGAGCGGCGGAGAGCTTCTCTGAAAGCAAGTGAAGTCAAGTCCG 732

RESULT 2
US-09-733-167A-2
; Sequence 2, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1042)..(1042)
; OTHER INFORMATION: n represents any one of a, c, t, and g.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1043)..(1043)
; OTHER INFORMATION: n represents any one of a, c, t, and g.
; NAME/KEY: misc feature
; LOCATION: (1114)..(1114)
; OTHER INFORMATION: n represents any one of a, c, t, and g.
US-09-733-167A-2

Query Match 21.3%; Score 193.8; DB 4; Length 1142;

Best Local Similarity 52.9%; Pred. No. 1.7e-34;
Matches 467; Conservative 0; Mismatches 407; Indels 9; Gaps 2;
QY 27 CCCGTGCTGAGAGAGAGATGATGCTGCTGACTACTACCGGATGCTGCTTCAACCGTAT 86
DB 54 CCAAGTGTGCGCAAGAGAGATGTGAGCAGAGAAACGGGGCTGTACACCTGACCGGAT 113
QY 87 GTTCAGAGTGTGAGCGGAGCACTGACCGAGTGCAGAGTCTGAGAGCTCTGAGCTTTCTGCT 146
DB 114 GTTCAGAGTGTGAGCGGAGCACTGACCGAGATGTGAGAGTGTGAGCTTTCTTCTCTCT 173
QY 147 GGATGAGCTCTGAGCGGCGGAGAGGCTTACCGCGGCGGAGGCGGCTAGAGCTCT 206
DB 174 GTTCAGAGTGTGAGCGGAGAGAGAGAGAGAGTCTAT---CGAATGAGAGCTGACTTCTT 230
QY 207 GCTGAGCTGAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 266
DB 231 AATTGAG 290
QY 267 CCTGCGGCTGCTGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326
DB 291 GCTGCGGAG 350
QY 327 AGTGTCTCAG 386
DB 351 GTGTGCTGCTGAG 410
QY 387 CTGCGGCTGAG 446
DB 411 CAG 464
QY 447 AGGCTCCCCCAAG 506
DB 465 TCCCACTATCTGAG 524
QY 507 ACGGCGGAG 566
DB 525 AGCCGAG 584
QY 567 CTCTGAG 626
DB 585 TCCCAAG 644
QY 627 TGGGCGAGCTTGAAG 686
DB 645 TGAGAGCTCTGAG 704
QY 687 GAGAGCTTGTGAG 746
DB 705 TGAGGCTTTAAACAG 764
QY 747 TGACATCAAGTCTGAG 806
DB 765 TGACATCAAGTCTGAG 824
QY 807 CGCCCTGCTGAG 866
DB 825 CTCTTATTAAG 884
QY 867 CCGGAGAGCTGTGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 909
DB 885 CCATGAG 927

RESULT 3
US-08-859-167-3
; Sequence 3, Application US/08859167
; Patent No. 6037461
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emed S.
; APPLICANT: Alnemri, Teresa
; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS

TITLE OF INVENTION: OF MAKING THE SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS: 17
ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,167
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1045 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 88..1044
US-08-859-167-3

Query Match 21.1%; Score 192.2; DB 3; Length 1045;
Best Local Similarity 52.8%; Pred. No. 3.8e-34;
Matches 466; Conservative 0; Mismatches 408; Indels 9; Gaps 2;

QY 27 CCCGCTGCGAGAGAGTAGTGCCTGGACTACTACGGGATGCTGCTTACCGTAT 86
DB 114 CCAGGTGTGGCCAGAGAGAGTAGAGACAGAACATGGCTGTACACCTGACCGCAT 173
QY 87 GTTGAGGTGTGGCGGGGCAATGACCGAGTGGAGCTGAGCTCCCTGGCTTTCTCT 146
DB 174 GTTGAACATCGTGGGCACTCATGTGACACACAGAGATGTGGCTTTCTTCTCTT 233
QY 147 GGATGAGGCTCTGGCGCGCGGAGGCTTAGCCGGGCGCGACAGCGGCTAGAGCTCT 206
DB 234 TGTGTATGTCAATATGACACAGAGCTGTGACTC---ATCGAATGACGCTGACTTT 290
QY 207 GCTGAGCTGAGCGCGCGGCGAGTGTGGCGAGAGCAACTGTGGCTGTGGGCAACT 266
DB 291 ATTGGCACTGGAGCGCGCGGCGGCTGTGATGAAGTAACCTTGGCCAGGTGTGCACT 350
QY 267 CTTGCGCTGTGTGGCGCGCGGCGAGACTTGTCCGACCTGTGGCGCGCAAGCGGCGGCG 326
DB 351 GCTGCGATCATCACTGCGACAGACTGTGCTCCCTACCTCACTCAAGAGAGAGCGGCG 410
QY 327 AGTGTCTCAGAGAGCTATAGCTATGAGCACTCCAGCTCTTCAAGAGAGAGAGAGGTAG 386
DB 411 TGTGTGCTGTATCTTGTAGACAGATATCTGAGAGAGACATCAATTCGCTATGTGACCC 470
QY 387 CTGCGGTGCGCGTGGCGAGTCAAGAGATTTTCAAGAGAGGTCAAGGTGAGAGAC 446
DB 471 CAGAGCCTCTAGATTCAGAGACCAAGAGCC-----TCCGAGCCCTTAAGACAGTGC 524
QY 447 AGGCTCCCCCAACCAAGCGGCGGAGAGTGTGGGCGGCGGCAATGTGTGTGCGAG 506
DB 525 TCCCACTATCTGTGTGTGTGTGCCCACTTGGGCTCTCAAGATGTGTAGCAAGCGGCG 584

QY 507 ACGGCGGCGAGAGGCGGCCCGGACCCGACAGCAGTCAAGCCCGCAGACCTTC 566
DB 585 AGCCCGAGGAGAGCCACACTTGGAGCCAGCGAAGCCGGAAGTCAAGACACAGA 644
QY 567 CTCTGAAGGCAATGACCTGTGACATCCGGCTCCGGGTTTGAGACAGTCTTCGAGCA 626
DB 645 TCCCAAGAGAGAGAGCATGTGACATCAAGCTCGGGGTTTGAGTAATCTCCAGCA 704
QY 627 TGGGCGAGCTTGGAGAGAGGCGTGGCATCCCGGCGCGCCAGCGCTGCGCGGCACT 686
DB 705 TGAGACTGTCTGAGAGGCAATGTCTTCTCTTAACAGAGAGACCACTTGAGCCCACT 764
QY 687 GGAAGTGTGGGAGGCGCAACGAGTGTGCGCTCAAGGAGACCTGTGCTGTGTTG 746
DB 765 TGAGCGCTTTAACAGGCGCAACCATTCCTCAAGTCCCGGAGCTGGGCTCATCATCT 824
QY 747 TGACATCAAGTCTCAAGAGCTCTCTATCTGAGAGCCCTTGTGGGCGCACTACCTGAGTGG 806
DB 825 TGACATCAAGTCTCTCAAGCTCACTACCTGATGACATTCGGGCTGATCATCATGAG 884
QY 807 CGCCGTGCTGAGGCGCTGCGGCGGCGTGTCTGACTGAGGCGCTGCGAGAGGCTGTGG 866
DB 885 CTCTTTATAGAGGCACTTAAGGTGTCTTATATACAGACTCCCTCAAGCAAGCTGTGG 944
QY 867 CCGGAGGCTGTGCTGCTGTGATCAAGTGTGATGAGGCTGAC 909
DB 945 CCATGAAGCATCAAGCTGTGTGTAAATGTAGAGAGAGAGAC 987

RESULT 4
US-09-109-273-3
Sequence 3, Application US/09109273
Patent No. 6063760
GENERAL INFORMATION:
APPLICANT: Alnemuri, Emad S.
APPLICANT: Fernandez-Alnemuri, Teresa
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/109,273
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1045 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: CDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 88..1044
US-09-109-273-3

Query Match 21.1%; Score 192.2; DB 3; Length 1045;
Best Local Similarity 52.8%; Pred. No. 3.8e-34;
Matches 466; Conservative 0; Mismatches 408; Indels 9; Gaps 2;

```

Qy 27 CCGGTGCTGGAGAGAGATGATGCTGAGCTAGCTAGCGGATGCTGCTTCAACCGTAT 86
Db 114 CCAAGTGTGCGCAGAAAGATGATGAGCAAGAACTGGCTGTACAGCTTGACCGCAT 173
Qy 87 GTTGAGGTGTGGGCGGCACTGACCGAGTGCAGCTGAGCTCTGCTCTTCTGCT 146
Db 174 GTTTGACATGTGGGCACTGATCTGACACAGAGATGTGCGGCTCTTCTTCTCTT 233
Qy 147 GATGAGGCTCTCGGCGCGCGGAGGCTTAGCCCGGCGCGAGCGGCTTAGAGCTCT 206
Db 234 TGTGTATGTATTGATATGACACAGCGTGACTC--ATCCGAATGAGAGTGACTTCTT 290
Qy 207 GCTGAGCTGAGAGCGCGCGGCGAGTGCAGAGAGCACTTGCGCTGCTGGGCACT 266
Db 291 ATTGGCACTGAGAGCGCGCGCGCTGTGATGAAGTAATCTTTCAGAGGTGTGCACT 350
Qy 267 CTTGCGGTGCTGCGCGCGCACTGCTGCGGCACTTGCGCGCAAGCGGCGCGCC 326
Db 351 GCTGCGCATATCACTGCGCAGCACTGCTGCGCTTACCTCAACCTCAAGAGAGACGGGC 410
Qy 327 AGTGTCTCCAGAACGCTATGCTATGACCTTCCAGCTCTTCAAGAGAGACAGGGTAG 386
Db 411 TGTGTGCTGTGATCTTGTAGCAAGATATCTGAGAGAGACATCAATTCGTATGTGACCC 470
Qy 387 CTCGCGTGTGCGCGTGCAGCTCAAGCAATTCGAAATTCAGCAGAGGTGAGTGGAGAC 446
Db 471 CAGAGCGCTCAGTATATCCAGAACCAAGGCG-----TCCGAGCGCTTAAACAGTGCC 524
Qy 447 AGGCTCCCCCAACCAAGCGGCGAGCGGAGTGTGGGCGCGCGCACTGTGTGTGCGAG 506
Db 525 TCCCACTATCTGTGTGTGTGTGCTCCCACTTCGGGTCTCAGATGTGTGAGCAAGCGGC 584
Qy 507 ACGGCGGCGGAGAGGCGCGCGCGAGCGGCAAGCGGCAAGTCAAGAGCGCGCGACCTTC 566
Db 585 AGCCCGGAGAGAGCCCACTTGGAGCGCGGAAAGCGCGGAGTCAAGTCAAGCAAGCA 644
Qy 567 CTCTGAAGCAAGTGAAGTGAATCCGAGTCCGAGTCCGAGTTCGAGCAAGTATGCGAGCA 626
Db 645 TCCCAAGGAGAGAGACATGTGATCAATCAAGCTGCGGCTTCCGCTGAATATGCGAGCA 704
Qy 627 TGGGCGAGCTTGGAGAGGCGGTGCAATCCGCGCGCGCGCGCGCTGCGCGCGAGCT 686
Db 705 TGAAGCTGTCTGTGAGGCGCAATGTCTTCAACAGCAGAGCCCACTTGAAGCGCGAGTT 764
Qy 687 GGAAGTGTGTGGAGAGCGCAAGCGAGTGTGCGGCTCAAGGAGCTTGGGCTCTGTGTTTG 746
Db 765 TGAAGCTGTTAACAGAGCGCAACCAATCTCAAGTCCCGGAGCTTGGGCTTCAATCACTTG 824
Qy 747 TGAATCAAGTGTCAAGAGCTCTCTATCTGAGAGCTTCTGAGGAGCACTTCACTGAGTGG 806
Db 825 TGAATCAAGTGTCTGAGAGCTCACTTCACTGAGAGCTTCTGAGGAGCACTTCACTGAGTGG 884
Qy 807 CGGCTGTCTGAGAGCGCTGCGGCGGTGTCTGATGAGAGCGCTTCCGAGAGGCTGTGGG 866
Db 885 CTCTTATTGAGAGCACTTAAAGGTGTCTTCACTCAAGAGCTTCCCTCAAGCAAGCTGTGGG 944
Qy 867 CCGGAGAGCTGTGTGCTGTGCTGCTGAGTGTGAGTGTGAGAGCTGAGC 909
Db 945 CCATGAAGCATCAAGCTGTGTGAATGTAGAGAGAGAGAGC 987

```

RESULT 5
US-09-276-993-3
Sequence 3, Application US/09276993

```

; Patent No. 6207801
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
; TITLE OF INVENTION: OF MAKING THE SAME
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6207801rlis
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,993
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/859,167
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TDU-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ. ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..1044
; US-09-276-993-3

```

Query Match 21.1%; Score 192.2; DB 3; Length 1045;
Best Local Similarity 52.8%; Pred. No. 3.8e-34;
Matches 466; Conservative 0; Mismatches 408; Indels 9; Gaps 2;

```

Qy 27 CCGGTGCTGGAGAGAGATGATGCTGAGCTAGCTAGCGGATGCTGCTTCAACCGTAT 86
Db 114 CCAAGTGTGCGCAGAAAGATGATGAGCAAGAACTGGCTGTACAGCTTGACCGCAT 173
Qy 87 GTTGAGGTGTGGGCGGCACTGACCGAGTGCAGCTGAGCTCTGCTCTTCTGCT 146
Db 174 GTTTGACATGTGGGCACTGATCTGACACAGAGATGTGCGGCTCTTCTTCTCTT 233
Qy 147 GATGAGGCTCTCGGCGCGCGGAGGCTTAGCCCGGCGCGAGCGGCTTAGAGCTCT 206
Db 234 TGTGTATGTATTGATATGACACAGCGTGACTC--ATCCGAATGAGAGTGACTTCTT 290
Qy 207 GCTGAGCTGAGAGCGCGCGGCGAGTGCAGAGAGCACTTGCGCTGCTGGGCACT 266
Db 291 ATTGGCACTGAGAGCGCGCGCGCTGTGATGAAGTAATCTTTCAGAGGTGTGCACT 350
Qy 267 CTTGCGGTGCTGCGCGCGCACTGCTGCGGCACTTGCGCGCAAGCGGCGCGCC 326
Db 351 GCTGCGCATATCACTGCGCAGCACTGCTGCGCTTACCTCAACCTCAAGAGAGACGGGC 410
Qy 327 AGTGTCTCCAGAACGCTATGCTATGACCTTCCAGCTCTTCAAGAGAGACAGGGTAG 386
Db 411 TGTGTGCTGTGATCTTGTAGCAAGATATCTGAGAGAGACATCAATTCGTATGTGACCC 470

```

Qy		387	CGGCGTTCGCGGTGGCAGCTAAGACGTTCTGCAAATTTCCAGACAGGGTGACTGGGAGAC	446
Db		471	CAGAGCCCTCAGTAGTATCCAGAACCAGGCC-----TCGCCAGCCCTCTTAACAAGTGCC	524
Qy		447	AGGCTCCCCCAACCAAGCCGCGACGGCGGAGTTCGGGGCCGGCCAGATGGTGTGTCAG	506
Db		525	TCCCACATATCTGTGTGTGTGTGCCCCACATTCCGGGTCTCAGATGTGTAGCAAGCGGCC	584
Qy		507	ACGGCGGCGAGAGAGGGGCCCGCAGCCGACCCACAGACAGTCAGAGCCCGCAGACCTTC	566
Db		585	AGCCCGAGGAGAGCACACTTTGGGAGCCAGCGAAAACCGCGAAAGTCAGTGAACCCAGA	644
Qy		567	CTCTGAAGGCAAAAGTACCTGTGACATTCGGCTCCGGTTTGAGCAGAGTACTGCGAGCA	626
Db		645	TCCCAAGGAGAGGACGACATGTGACATCAGCTCGGGTGTGGGCTGTGATCTCCAGCA	704
Qy		627	TGGGCGAGCCTTGGAGCAGCGGCGTGGCATCCCGGGCGGCCAGGCGCTGCGCGGCGACT	686
Db		705	TGAACATCTCTGACAGGGCAAATGCTTCTCTMAACAGAGGAGCCACATTGAGCCCGCTT	764
Qy		687	GGAAGTGTGGGAGAGGCCACCGCAGTGTGCTCGCTCAAGGGAGCTTGAGCTGTGTGTTG	746
Db		765	TGAGCGCTTTAACAGAGCCCAACACATCTCAAGTCCCGGACCTTGCGCTCATCATCTG	824
Qy		747	TGACATCAAGTCTCAGAGCTCTCTATCTGSAAGCCTTCTGCGGCGCACTACCTAGTGG	806
Db		825	TGACATCAAGTCTCTGAGCTCACCTACCTGATGACATTTCTGGGTGACTATCAATATGG	884
Qy		807	CGCCCTGCTGAGGCCCTTCGGGGCGGTGTTCTTGACTGAGGCCCTTCGAGAGGCTGTGGG	866
Db		885	CTCTTATTAAAGGCACCTTAAAGGTGTCTTATCATCAGACATCCCTCAAGCAAGCTGTGGG	944
Qy		867	CCGGGAGGCTGTTCGCTGCTGCTCAGTGTGATGAGGCTGAC	909
Db		945	CCATGAGCCATCAAGCTGCTGTAAATGTAGACGAGGAGAC	987
 RESULT 6 US-09-723-450-3				
Sequence 3, Application US/09723450				
Patent No. 6576751				
GENERAL INFORMATION:				
APPLICANT: Alnemri, Emad S.				
TITLE OF INVENTION: Fadd-like Anti-Apoptotic Molecules, Methods Of Using The Same, An				
FILE REFERENCE: TUD2445				
CURRENT APPLICATION NUMBER: US/09/723,450				
PRIORITY FILING DATE: 2000-11-28				
PRIOR APPLICATION NUMBER: 09/276,993				
PRIOR FILING DATE: 1999-03-26				
PRIOR APPLICATION NUMBER: 08/859,167				
PRIOR FILING DATE: 1997-05-20				
NUMBER OF SEQ ID NOS: 17				
SOFTWARE: PatentIn version 3.0				
SEQ ID NO 3				
LENGTH: 1045				
TYPE: DNA				
ORGANISM: Artificial Sequence				
FEATURE:				
NAME/KEY: misc feature				
OTHER INFORMATION: No. 6576751el Sequence				
NAME/KEY: CDS				
LOCATION: (88) ..(1044)				
US-09-723-450-3				

	Query Match	Similarity	Score	DB	Length
Best local	52.8%	Pred. No.	3.8e-34		
Matches	466	Conservative	0	Mismatches	408
				Indels	9
				Gaps	2
Qy	27	CCCGTGTGGAGAGAGATGATGCTCTGCACTACGAGATGCTGTGCGTTCAACCGAT	86		
Db	114	CCAGGTGTGGCCAGAGAGCATGTGAGACAGAGACATGTGGCGTGTACAGCTTCGACCGAT	173		

OY	87	GTTGAGAGTGTGGGCGGGCAA	CTGACCGAGTGGAGCTGGAGCTCCGCGCTTTCGCT	146	
Db	174	GTTTGACATCGTGGGCACTAC	TTCTGACACACAGAAATGTGGCGTGCTTTTCTCTCT	233	
OY	147	GGATGAGGCTTCTGGGCGCG	CCGAGAGGCTTAGGCCCGGCGCCGACGGGCTTAGAGCTCT	206	
Db	234	TGTTGATGTACTTATATAC	CAACGAGCGTGGACCTC---ATCCGAATATGAGCGTGACTTCTT	290	
OY	207	GCTGGAGCTGGAGCGCGCG	GGGAGTGGGGGAGAGACA	CTTGCGGCTGCTGGGGCACT	266
Db	291	ATTGGCACTGGAGGGCCAG	GGGCGCCCTGTGATGAATAATCTTTGGCCAGGTGTCTGCACT	350	
OY	267	CCTGGGGGTGTGGGCGGCA	CGACCTGTGCGCACCTGGCGGCAAGCGGCGCGCGCC	326	
Db	351	GCTGGGCAATCACTGGCA	CGACCTGTGCGCACCTTAGCTACCTCAAGAGAGAGCGGCG	410	
OY	327	AGTGTCTTCAGAACGCTAT	AGCTTAGGCACTTCAGCTCTTCAAGAAGSACAGAGGTTAG	386	
Db	411	TGTGTGCCCTGATCTTGT	ATGACAGATATCTGGAGAGACATCAATTTGGCTATGTGACCCC	470	
OY	387	CTGCGGTGCGCGTGGCAG	TTCAGACATTTCTGCAAAATTTCTACGACGGTCACTGGGAGAC	446	
Db	471	CAGAGCCCTCAGTATCC	AGAACCAAGGCC-----TCCCGAGCCCTTAAGAACATGCGC	524	
OY	447	AGGCTCCCCCACAACGA	CGGCGGAGTCTGGGGCGCGGCCCAAGTGTGTGGTCCAG	506	
Db	525	TCCCCACTATCTTGTGT	GTGTGTGTGTGCCCACTTGGGCTCTCAGATGTGTAGCAACGGCC	584	
OY	507	ACGGCGGCGAGAGAGGG	CGCCGACGCCGACAGCATGACAGCCCGGCAAGCTTTC	566	
Db	585	AGCCCGAGGAGGAGCA	CCACACTTTGGGAGCGACGAAACGCGGAAATGACGACACCGA	644	
OY	567	CTCTGAAGGCAAAATGA	CACTGTGACATTCGCGCTCGGGTTTCGACAGAGTACATGCGAGCA	626	
Db	645	TCCCAAGAGAGAGAC	AGATGTGACATGACTCGGGTGTGGGCTGAAATCTGCAAGCA	704	
OY	627	TGGGCAAGCTTTGGAGA	GAGGCTGTGAGATCCCGGCGGCGCCGACGCGCTGCGCGGACGT	686	
Db	705	TGAGACGTCTCTGAG	GGGCAATGCTTCTCTTAACAAGAGGACCACTTGAAGCGCGAGTT	764	
OY	687	GGAAGTGTGGGAGAG	CGCACCGCAGTGTGCGCTCAAGGSACTTGCGGCTCTGTGTGTTTG	746	
Db	765	TGAGCGCTTTTAAC	CGAGGCAACACATCTCTCAATCCCGGGACCTGGGCTTCATCATCTCG	824	
OY	747	TGACATCAAGTTCTCAG	AGCTCTCTTAATCTGGAAGCGCTTCTGGGGCGGACTACCTGAGTGG	806	
Db	825	TGACATCAAGTCTCTG	AGCTCACTTACCTTGATGCAATTTCTGGGCTGTACTATCAATCATATGG	884	
OY	807	CGCCCTGTGACAG	CGCTTGGCGGGCGTGTCTTGACTAGAGCCCTGCGACGAGAGGCTGTGGG	866	
Db	885	CTCTTTATTAAG	GCACCTTAAGAGGTGCTTATCAACAGACTCTGTCGG	944	
OY	867	CCGGAGGCTGTTCCT	GCTGTGTCAGTGTGAGATGAGGCTGAC	909	
Db	945	CCATGAAGCCATCAAG	CTGCTGTGTAAATGTAGACGAGAGGAC	987	

```

RESULT 7
US-09-733-167A-4
; Sequence 4, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733, 167A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547
; PRIOR FILING DATE: 1998-06-08
198 25 621.3

```

NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1142
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-733-167A-4

Query Match 21.0%; Score 190.6; DB 4; Length 1142;
 Best Local Similarity 52.7%; Pred. No. 8.9e-34;
 Matches 465; Conservative 0; Mismatches 409; Indels 9; Gaps 2;

27 CCGGTGCTGGAGAGAGATAGTCCCTGAGCTACTAGAGATGTCGCTTACCCGTAT 86
 54 CCAAGTGTGCGCCGAAAGAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 113
 87 GTTCGAGAGTGTGGGCGGGGCACTGACCGAGTGGAGAGAGAGAGAGAGAGAGAGAG 146
 114 GTTCGAGAGTGTGGGCGGGGCACTGACCGAGTGGAGAGAGAGAGAGAGAGAGAGAG 173
 147 GATGAGAGTCTTGGCGCGCGCGAGGCTTACCGCGGCGCGAGCGGCGCTAGAGCTCT 206
 174 TGTGATGTTATTGATGATGACATGAAAG--TGACCTCATCGAAGATGAGAGTGTCTT 230
 207 GCTGAGAGTGGAG 266
 231 ATTGGACATGAG 290
 267 CTTGCGGCTGCTGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326
 291 GCTGGCATCATCATCTGCGCATGAGCTTGTGCTGCTTACCTTCAAGAGAGAGAGAG 350
 327 AGTGTCTCCAG 386
 351 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 410
 387 CTGCGGCTGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446
 411 CAGAGCCCTAGTGAAG 464
 447 AGGCTCCCCCAACCAAG 506
 465 TCCCACTATCTGT 524
 507 ACAGGCGGAG 566
 525 AGCCCGAG 584
 567 CTCTGAG 626
 585 CCGGAG 644
 627 TGGGCGAG 686
 645 TGAG 704
 687 GAG 746
 705 TGAG 764
 747 TGACATGAG 806
 765 TGACATGAG 824
 807 CGGCGCTGCTGAG 866
 825 CTGATTTATGAG 884
 867 CCGGAG 909
 885 CATGAG 927

RESULT 8

US-08-859-167-5

Sequence 5, Application US/08859167

Patent No. 6037461

GENERAL INFORMATION:

APPLICANT: Alnemri, Emad S.

TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF

TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s

STREET: One Liberty Place, 46th floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: WINDOWS

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/859,167

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3439

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1200 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 32..988

US-08-859-167-5

Query Match 21.0%; Score 190.6; DB 3; Length 1200;
 Best Local Similarity 52.7%; Pred. No. 9e-34;
 Matches 465; Conservative 0; Mismatches 409; Indels 9; Gaps 2;

27 CCGGTGCTGGAGAGAGATAGTCCCTGAGCTACTAGAGATGTCGCTTACCCGTAT 86
 58 CCAAGTGTGCGCCGAAAGAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 117
 87 GTTCGAGAGTGTGGGCGGGGCACTGACCGAGTGGAGAGAGAGAGAGAGAGAGAGAG 146
 118 GTTCGAGAGTGTGGGCGGGGCACTGACCGAGTGGAGAGAGAGAGAGAGAGAGAGAG 177
 147 GATGAGAGTCTTGGCGCGCGCGAGGCTTACCGCGGCGCGAGCGGCGCTAGAGCTCT 206
 178 TGTGATGTTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 234
 207 GCTGAGAGTGGAG 266
 235 ATTGGACATGAG 294
 267 CTTGCGGCTGCTGGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326
 295 GCTGGCATCATCATCTGCGCATGAGCTTGTGCTGCTTACCTTCAAGAGAGAGAGAG 354
 327 AGTGTCTCCAG 386
 355 TGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 414

Oy	387	CTGCGGTGGCCGGTGGGACAGTCAAGCAGTTTGTGAAATTCTCAGCAGGGTCAATGTGGAGAC	444
Db	415	CAGAGCCCTCAGTAGACCCAGAACCGAGGCC-----TCCCAACCCCTCTAATAACAGTGCC	468
Oy	447	AGGCTTCCCCCAACCAACGAGCGGAGCGGCGAGTCCGGGCGCGGCCCACTGATGGTGGCCAG	500
Db	469	TCCCACTATCTCTGTGTGTGTGCTGCCCCCACTTGGGTTCTCAATGTGTAGTAAGCGGCC	528
Oy	507	ACGGCGGCGGAGAGGGGGCCCCAGCCGCAACCCAGCAGCACTAGAGCCCGCCAGACCTTC	566
Db	529	AGCCGAGGAGAACCAACACTTGTGGAGCGACGAAACCGCGGAAGTGGTGAACACAGA	588
Oy	567	CTCTGAAGCCAAAGTATCCTGTGACATCCGGCTCCGGGTTCGAGCAGAGTACTGCGAGCA	628
Db	589	CCCGAAGGAAAAAGCAGACATGTGATATCAGGCTCCGAGTTCCGGCGGAATACTGCCAGCA	648
Oy	627	TGGGCGAGCTTGTGAGCAGAGGGGTGGCATCCCGCGCGGCCAGGCGTGGCGGGCGAGCT	686
Db	649	TGAGACGGTCTTGCAAGGSCAATGTCTTTCATTAAGCAGAGACCACTTAGGCGCCAGTT	708
Oy	687	GGACCTGTTTGGGAGGCGACCGCAGTGTGCTGCCTCAAGGGACCTTGAGGCTCTGTGGTTTG	746
Db	709	TGAGGCGCTTTAAACAGAGCCCAACACTATCTCTCAAGTCCCGGAGACTTGAGGCTCATCATCTG	768
Oy	747	TGACATCAAGTTCTCAGAGCTCTCCTATCTGACAGCGCTTCTGTGGCGCATCACTCGTAGCTGG	806
Db	769	TGACATCAAGTTCTCTGAGCTCACTCAACCTCGACGATTTCTGGCGAGACTCAATTATGG	828
Oy	807	CGCCCTGTGCGAGGCGCTTGCGGGGCGTGTCTCTGACTGAGGCCCTTGCGAGAGGCTGTGGG	866
Db	829	CTCATATTATGAGGCGCACTGAAGGTGTCTTCATCAACAGACTCTCTCAAGCAACTGTGGG	888
Oy	867	CCGGGAGAGCTGTGCGCTGTGCTGTGACTGTGATGTAAGGCTGAC	909
Db	889	CCATGAAGCCATCAAGCTGTGGTGTGACGTGATGGATGGAGAGAC	931

```

1 RESULT 9
2 US-09-109-273-5
3 : Sequence 5, Application US/09109273
4 : Patent No. 6063760
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Alnemri, Enad S.
9 : APPLICANT: Fernandez-Alnemri, Teresa
10 : TITLE OF INVENTION: PADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
11 : TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
12 : TITLE OF INVENTION: OF MAKING THE SAME
13 : NUMBER OF SEQUENCES: 17
14 :
15 : CORRESPONDENCE ADDRESSES:
16 : ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760r1s
17 : STREET: One Liberty Place, 46th floor
18 : CITY: Philadelphia
19 : STATE: PA
20 :
21 : COUNTRY: USA
22 :
23 : ZIP: 19103
24 :
25 : COMPUTER READABLE FORM:
26 : MEDIUM TYPE: Floppy disk
27 : COMPUTER: IBM PC compatible
28 : OPERATING SYSTEM: WINDOWS
29 : SOFTWARE: WordPerfect
30 :
31 : CURRENT APPLICATION DATA:
32 : APPLICATION NUMBER: US/09/109,273
33 : FILING DATE:
34 :
35 : CLASSIFICATION:
36 : PRIOR APPLICATION DATA:
37 : APPLICATION NUMBER: 08/859,167
38 : FILING DATE:
39 :
40 : ATTORNEY/AGENT INFORMATION:
41 : NAME: Deluca, Mark
42 : REGISTRATION NUMBER: 33,229
43 : REFERENCE/DOCKET NUMBER: TJU-
44 : TELECOMMUNICATION INFORMATION:

```

```

? TELEPHONE: (215) 568-3100
? TELEFAX: (215) 568-3439
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1200 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: both
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 32..988
US-09-109-273-5

```

US-09-109-273-5

Query Match	21.0%;	Score 190.6;	DB 3;	Length 1200;
Best Local Similarity	52.7%;	Pred. No. 9e-34;		
Matches 465;	Conservative 0;	Mismatches 409;	Indels 9;	Gaps 2;

OY	27	CCCGGCTGAGGAGAGATGATGATGCTGAGATCACTACGAGGATGCTGCGCTTCAACCGTAT	86
Db	58	CCAGGTGTGCCCCGAAAGACGTGGGAGCAAGAAATAGGCGCTCTACAGCGCTCCACCGCAT	117
OY	87	GTTTCGAGGTGTGCGCGGGCAACTGACCGAGTGCAGCTGCAGCTCTGCGCTTCTGCT	146
Db	118	GTTTCGACATCGTGGGCACTCACTAAACACAGAGATGTCGAGTGCCTTCTCTCTTT	177
OY	147	GGATGAGGCTCCTGCGCGCGCCCGAGGCTTAGCCCGCGGCCGACCGGCTTGAGGCTCCT	206
Db	178	TGTTGATCTTATTATGATGACATGAACG--TGAGCTCATTCGAAATGACGTGACTTCTT	234
OY	207	GCTGAGCTGAGAGCGCCGCGGGCAGTGCAGCGAGAGCAACTGCGGCTCTGGGGCAACT	266
Db	235	ATTGGCACTGAGAGCGCCAGGGCCGCTGTGACGAGAGTAATCTTCCGCGAGTCTCGAGCT	294
OY	267	CCTGCGCGTGTGCGCCGCGCAAGACTGTGCAGCACTGCGCGCAAGCGCGCCGCGC	326
Db	295	GCTGCGCATCATCACTCCGCCATGACTTGCTCGCCCTACGTTACTCTCAAGAAAGACGAC	354
OY	327	AGTGTCTCCAAACGCTATACCTATAGGCACTCCAGCTCTTCAAGAGAGACAGAGGTAG	386
Db	355	TGTGTGCCCTATCTTTGAGCAAACTATCTGGAGGAAACATCAATTCGCTATGTGACCC	414
OY	387	CTGCGGTGCGCGTGCAGTCAAGCAGTTCTGCAAAATCTCAGCAGGTCAGTGGAGAG	446
Db	415	CAGAGCCCTCAGTACCCAGAACGAGGCGC-----TCCCGACCCCTCTAAAACAGTGC	468
OY	447	AGGCTCCCCCAACCAAGCGGAGCGGCGGAGTGGGGCGCGGCCCACTGTGTGTCCAG	506
Db	469	TCCCCACTATCTGTGTGTGTGCTGCGCCCACTTCGCGGTTCTCAATGTGTAAGCGGCG	528
OY	507	AACGCGCGGAGAGAGGGGCCCAAGCCGACCCAGACAGTACAGACCCGCGACGACTTC	566
Db	529	AGCCGAGGAGAACCAACACTTTGGAGGCAAGCAAGAAAGCGCGGAAGTCGTACACACAG	588
OY	567	CTCTGAAGGCAAGTGAACCTGTGACATCCGGCTCGGGTCTCAGAGAGTACTGCGAGCA	626
Db	589	CCCGAAGAAAGACGACATGTGATATAGGCTCCGAGTTGGGGCGGAATCTGCGACGA	648
OY	627	TGGGCGAGCTTGGAGCAGGCGGTGACATCCCGCGGCGCCAGCGCTGCGCGGCGAGCT	686
Db	649	TGAGACGGCTCTGCAAGGCAATGTCTTCTCCAAATAGAGAGACCACTTGAGCGCAGATT	708
OY	687	GGAGGTGTTTGGGCAAGCCACCGCAGTGTGTGCGTCAAGGGAACCTGGGCTCTGTGTTTG	746
Db	709	TGAGCGCTTTAAACAGAGCCAAACATATCTCAAGTCCCGGAGCTGTGGGCTTCATCATCTG	768
OY	747	TGACATCAAGTTCTCAGAGCTCTCCTATCTGGACGCTTCTGGGCGGACCTACCTGAGTGG	806
Db	769	TGACATCAAGTTCTCTGAGCTCACCCTACCTCGAGCCGACTTCTGGCGAGACTAATTATGG	828
OY	807	CGCCCTGCTGAGGCGCTGCGGGCGTGTCTCTGACTGAGGCGCTTGCAGAGAGGCTGTGG	866
Db	829	CTCATTTATTAAGGCACTGAAGAGGTGCTTATCATCAACACTCTCTCAAGCAAGCTGTGGG	888

US-09-723-450-5

Query Match 21.0%; Score 190.6; DB 4; Length 1200;
Best Local Similarity 52.7%; Pred. No. 9e-34;
Matches 465; Conservative 0; Mismatches 409; Indels 9; Gaps 2;

QY 27 CCGCTGCTGGAGAGAGATGAGTCCCTGGAAGTCTTACCGAGATGCTGCTTCAACCGTAT 86
DB 58 CCAAGGTGGCCCGAAGAGCGTGGGAGCAAGAACATGGCTTACAGCTCCACCGCAT 117
QY 87 GTTCCGAGGTGGTGGCCGGCACTGACCGAGTGGAGTGGAGTCTCTGGGCTTTTGTCT 146
DB 118 GTTCCGAGGTGGTGGCCCGCACTTACACAGAGATGTCGAGTGTCTTCTTCTTCTT 177
QY 147 GATGAGAGTCTCTGCGCCCGCCGAGGCTTACCGCGGCGCCGAGCGGCTTACGAGTCTT 206
DB 178 TGTGTAAGTATTATATATACATGAAAG---TGACTATCTCGAAATGAGCTGACTTCTT 234
QY 207 GCTGAGCTGAGAGCGCCGCGGAGTGGCGGAGAGCAACCTGCGCTGCTGGGAGCACT 266
DB 235 ATTGGCACTGGAGCGCCAGGGCCGCTGTGACGAGAGTAACTTTGCGCAGGTGTGCAAGCT 294
QY 267 CTTGGCGGTGTGGCCCGCCGACGACCTGCTCCGACCTGGCGCGCAAGCGGCGCCGCG 326
DB 295 GCTGGGATATATCACTGCGCACTGCTGCTGCTTACGTTACTCTCAAGAGAGACGAGC 354
QY 327 AGTGTCTCCAGAAAGCTATAGCTATGAGCACTCCAGCTCTTCAAGAGAGACGAGGGTAT 386
DB 355 TGTGTGCTCTGATTTTGTAGCAAGTATCTGAGAGAAACATCAATTGCTATGTATCCCC 414
QY 387 CTGCGGTGCGCGGTGGAGAGTCAAGCACTTTCGAAATTTCTGAGAGGGTCACTGAGGAG 446
DB 415 CAGAGCCTCAGTACCCAGAACGAGAGC-----TCCCCAGCCTCTTAAACAGTGGCC 468
QY 447 AGGTCTCCCCCAACCAAGCGGAGCGGCGGAGTGGGAGCGGCGCCAGTGTGTGCCAG 506
DB 469 TCCCACTATCTGT 528
QY 507 ACCGCGCGGAGAGAGGCGCCGAGCGGACCCGAGAGAGTCAAGAGCGGCGGAGAGCTTC 566
DB 529 AGCCCGAGGAGAGACACACTTTGGAGCGAGCGAAAGCGCGAGTGTGTGTGTGTGTGTGT 588
QY 567 CTCTGAGAGCAAGTGAAGTGTGATCTGAGCTCGGCTCGGCTTTCGAGCAAGTATCTGAG 626
DB 589 CCGGAGAGAAAGAGACATGTGATATCGAGCTCCGAGTTCGGGCGGAAATCTGCGAGCA 648
QY 627 TGGGCGAGCTTGGAGAGAGGCGGTGCACTCCGCGCGGCGCCAGCGCTGCGCGGAGCT 686
DB 649 TGAACCGGCTCTGCAAGGCAATGTCTTCCATTAAGGAGAGACCACTTGAAGCGCAGTT 708
QY 687 GAGCTGTGTGGGAGGCAACCGGAGTGTGCGCTCAAGGAGAGCTGGGCTGTGTGTGTG 746
DB 709 TGAAGCTTTTAAACAGGCGCAACATATCTCAAGTCCCGGAGCTGGGCTCATCATCTG 768
QY 747 TGAATCAAGTCTTCAAGAGCTCTCTATCTGAGAGCTTCTGGGCGCACTACCTGAGTGT 806
DB 769 TGAATCAAGTCTTCTGAGCTCACTACCTGAGAGCAATTCGGGAGAGCTATTAATGG 828
QY 807 CGCCCTGCTGAGGCGCTGCGGCGGCTGTCTTCACTGAGGCGCTTGGAGAGGCTGTGGG 866
DB 829 CTCATTATTAGAGCACTGAAGGCTGTCTTCACTACAGACTCTCTCAAGCAAGCTGTGGG 888
QY 867 CCGGAGAGGCTGTGGCGGCTGTCAAGTGTGATGAGGCTGAC 909
DB 889 CCAATGAAGCAATCAAGCTGTGTGAACGTGATGAGAGGAGAC 931

RESULT 12

US-09-733-167A-8
; Sequence 8, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus

APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; CURRENT FILING DATE: 2000-12-08/01712
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Truncated segment of the DNA sequence encoding human BADD, compr
; US-09-733-167A-8

Query Match 16.1%; Score 146.2; DB 4; Length 630;
Best Local Similarity 60.4%; Pred. No. 8e-24;
Matches 241; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 511 CCGCGAGAGAGGCGCCAGCGGACCCGAGAGAGTCAAGAGCGGCGGAGAGCTTCTCT 570
DB 178 CAGAGGAGAGGACACACTTTGGAGCGAGCAAAAGCGGAGAGTCAAGTACACAGATCCC 237
QY 571 GAAGGCAAGTGAAGCTTGTGACATCTGCGCTCCGGTGTGAGCAGATCTGAGAGATGG 630
DB 238 AAGGAGAGAGAGACATGTGACATCACTGCGGTTGGGCTGTAAATCTGCAAGATGAG 297
QY 631 CAGAGCTTGAAGAGAGGCGGTGAGATCCCGGCGGCGCCAGGCGCTGGGCGGAGCTGAG 690
DB 298 ACTGCTGTGAGGCAATGTCTTCTTAAAGAGAGAGCCACTTGAAGCGCACTTTGAG 357
QY 691 GTGTTGGGAGAGGCGGAGAGTGTGCTGCTCAAGGAGAGTGTGCTGTGTGTGTGAG 750
DB 358 CGCTTTAAACAGGCGCAACATCTCTCAAGTCCCGGAGCGGCTCATCATCTGTGAG 417
QY 751 ATCAAGTCTCAAGAGCTCTCTATCTGAGAGCTTCTGGGCGAGTACCTGAGTGGCGC 810
DB 418 ATCAAGTCTCTGAGCTCACTACCTGAGTCAATCTGCGGTGCTATCATCAATGCTCT 477
QY 811 CTGCTGAGGCGGCGGCGGCGGTCTCTGAGTCAAGGCGGCGGAGAGGCTGTGGGCGG 870
DB 478 TTATTAGAGGACTTAAAGGTGTCTTATCATCAAGACTCTTCAAGCAAGCTGTGGCCAT 537
QY 871 GAGGCTGTGGCGCTGTGCTGAGTGTGATGAGGCTGAC 909
DB 538 GAAGCATCAAGCTGTGTGAATGTAGAGAGGAGAGAC 576

RESULT 13

US-09-733-167A-7
; Sequence 7, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 342
; TYPE: DNA

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Truncated segment of the DNA sequence encoding human DEDD, compri
OTHER INFORMATION: sing nucleotides 28-369 of SEQ ID NO. 2.
US-09-733-167A-7

Query Match 9.7%; Score 88.4; DB 4; Length 342;
Best Local Similarity 57.3%; Pred. No. 7.5e-11;
Matches 180; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

QY 27 CCGGCTGGGAGAGAGATAGTCCGAGCTACTAGAGGATGCTGCTTACCGTAT 86
DB 27 CCGGCTGGGAGAGAGATAGTCCGAGCTACTAGAGGATGCTGCTTACCGTAT 86
QY 87 GTTGAAGTGTGGGAGGAGCACTGACCGAGTCCGAGCTGCTGCTTCTGCT 146
DB 87 GTTGAAGTGTGGGAGGAGCACTGACCGAGTCCGAGCTGCTGCTTCTGCT 146
QY 147 GATGAGGCTCCGCGGCGGCGGAGCTTACGCGGCGGCGGAGGCTTACGCTCT 206
DB 147 GATGAGGCTCCGCGGCGGCGGAGCTTACGCGGCGGCGGAGGCTTACGCTCT 203
QY 207 GCTGAGAGTGGAGCGCGCGGCGGAGTGGCGAGAGCACTGCGGCTGCTGGGCACT 266
DB 204 ATTGCACTGGAGCGCGGCGGCGGCTGTATGAAAGTCTTCCGAGGCTGCACT 263
QY 267 CCGGCGGCTGCTGCGCGGCGGAGCACTGCTGCGGAGCTGCGGCGGAGGCGGCGG 326
DB 264 GCTGCGGATCATCATCGCGGAGCACTGCTGCGGAGCTGCGGAGCTGCGGAGGCGG 323
QY 327 AGTGTCTCCGAGAC 340
DB 324 TGTGTGCTCCGATC 337

RESULT 14

US-07-704-288C-2
Sequence 2, Application US/07704288C
Patent No. 5399680
GENERAL INFORMATION:
APPLICANT: LAMB, CHRISTOPHER J.
APPLICANT: ZHU, QUN
TITLE OF INVENTION: PLANT DEFENSE GENES AND PLANT DEFENSE REGULATORY
TITLE OF INVENTION: ELEMENTS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER, BRUEGEMANN & CLARK
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: California
COUNTRY: United States
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/704,288C
FILING DATE: 22-MAY-1991
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P31 8899
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 546-4737
TELEFAX: (619) 546-9392
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1151 base pairs
TYPE: nucleic acid

STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 55..1062
US-07-704-288C-2

Query Match 5.6%; Score 50.8; DB 1; Length 1151;
Best Local Similarity 57.6%; Pred. No. 0.03;
Matches 91; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 159 TGGCGCGCGGAGGTTAGCCCGGCGGCGGAGGCGGCGGAGGCTGAGCTGAGGTTGA 218
DB 71 TGTGGGCAATGTGGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 130
QY 219 GCGCGCGGCGGAGGCGGCGGAGGCACTGCGGCGGCGGCGGCGGCGGCGGCGGCT 278
DB 131 GCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 190
QY 279 GCGCGCGGAGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 316
DB 191 GCTCCAGCTCCGAGCTACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 228

RESULT 15

US-08-093-372-1
Sequence 1, Application US/08093372
Patent No. 5530187
GENERAL INFORMATION:
APPLICANT: Lamb, Christopher J.
APPLICANT: Zhu, Qun
APPLICANT: Maher, Eileen A.
APPLICANT: Dixon, Richard A.
TITLE OF INVENTION: TRANSGENIC PLANTS CONTAINING MULTIPLE
TITLE OF INVENTION: DISEASE RESISTANCE GENES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 South Flower Street, Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: U.S.A.
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,372
FILING DATE: 16-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1151 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 55..1062
OTHER INFORMATION: /product= "RICE CHITINASE"
US-08-093-372-1

Query Match 5.6%; Score 50.8; DB 1; Length 1151;
 Best Local Similarity 57.6%; Pred. No. 0.03;
 Matches 91; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY	159	TGGGCGCGCGGAGGCTTAGCCCGGCGCGAGCGGCTAGAGCTCCTGCTGAGCTGA	218
Db	71	TGATGGCCATGTGTGCGCCAGCCCTTCTTCGCGCGCGCCGTGCATGCGAGCAGTGCGCA	130
QY	219	GCGCGCGGCGGAGTGTGCGCGAGAGCAACTGCGGCTGTGGGCAACTTCTGCGGCTGCT	278
Db	131	GCCAGGCGCGGCGGCGGTGTGCCCAACTGTCTGTGCAAGCAGTTGCGGCTGTGCG	190
QY	279	GCGCGCGCAGCACTGTGCTGCGCACTGTGCGCGCAAGC	316
Db	191	GCTCCACTCTCCGACTACTGTGCGCGCGGATGCCAGAGC	228

Search completed: March 23, 2005, 16:17:54
 Job time : 169.414 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 09:53:17 ; Search time 563.567 Seconds
(without alignments)
9612.048 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgtacccggtcgcac.....tcagtgatgagcgcac 909

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 5552208 seqs, 297965951 residues 11104416

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907.4	99.8	1230	16	US-10-296-539-4
2	907.4	99.8	1924	14	US-10-001-254-17
3	907.4	99.8	2045	14	US-10-013-477-2
4	907.4	99.8	2045	15	US-10-106-698-349
5	895.4	98.5	2044	9	US-09-925-302-315
6	895.4	98.5	2044	10	US-09-925-302-315
7	799.4	87.9	1966	9	US-09-823-830A-390
8	581.2	63.9	1570	17	US-10-286-115-481
9	541	59.5	1067	15	US-10-037-270-853
10	541	59.5	1067	17	US-10-117-722-853
11	539.2	59.3	1554	9	US-09-799-777-103

C	12	326.6	35.9	111084	18	US-10-723-860-1627	Sequence 1627, Ap
	13	301.4	33.2	303	14	US-10-001-254-7	Sequence 7, Appl
	14	193.8	21.3	1142	2	US-09-733-167-2	Sequence 2, Appl
	15	193.8	21.3	2261	9	US-09-920-300A-1684	Sequence 1684, Ap
	16	193.8	21.3	2261	13	US-10-033-528-1684	Sequence 1684, Ap
	17	193.8	21.3	2261	16	US-10-099-926-1684	Sequence 1684, Ap
	18	192.2	21.1	1045	9	US-09-935-223-3	Sequence 3, Appl
	19	192.2	21.1	2079	18	US-10-357-930-24750	Sequence 24750, A
	20	190.6	21.0	1142	9	US-09-733-167-4	Sequence 4, Appl
	21	190.6	21.0	1200	9	US-09-935-223-5	Sequence 5, Appl
	22	157.2	17.3	484	10	US-09-918-995-12268	Sequence 12268, A
	23	152.4	16.8	451	10	US-09-918-995-28421	Sequence 28421, A
	24	146.2	16.1	630	9	US-09-733-167-8	Sequence 8, Appl
	25	112	12.3	553	10	US-09-918-995-10904	Sequence 10904, A
	26	111	12.2	278	14	US-10-066-543-17	Sequence 17, Appl
	27	105.4	11.6	1107	9	US-09-785-651-104	Sequence 104, Ap
	28	88.4	9.7	342	9	US-09-733-167-7	Sequence 7, Appl
	29	52.8	5.8	2073	18	US-10-489-988-3	Sequence 3, Appl
	30	52.8	5.8	2328	16	US-10-114-153-57	Sequence 57, Appl
	31	52.8	5.8	2382	9	US-09-789-404-3	Sequence 3, Appl
	32	52.8	5.8	2511	18	US-10-489-988-1	Sequence 1, Appl
	33	52.8	5.8	2636	9	US-09-789-404-1	Sequence 1, Appl
	34	50.4	5.5	2584	18	US-10-437-963-42395	Sequence 42395, A
	35	50.2	5.5	1361	18	US-10-425-115-180284	Sequence 180284, A
	36	50.2	5.5	903	18	US-10-425-115-170686	Sequence 170686, A
	37	48.8	5.4	4020	10	US-09-796-679-5	Sequence 5, Appl
	38	48.8	5.4	137560	18	US-10-481-112-1	Sequence 1, Appl
	39	48.4	5.3	1943	16	US-10-146-733-14	Sequence 14, Appl
	40	48.4	5.3	1943	18	US-10-757-262-91	Sequence 91, Appl
	41	48.4	5.3	1943	18	US-10-768-158-43	Sequence 43, Appl
	42	48.4	5.3	2080	17	US-10-257-022-30	Sequence 30, Appl
	43	48.4	5.3	1366	18	US-10-437-963-2075	Sequence 2075, Ap
	44	47.2	5.2	417	15	US-10-156-761-3273	Sequence 3273, Ap
	45	47.2	5.2	9025608	15	US-10-156-761-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-10-296-539-4

Sequence 4, Application US/10296539

Publication No. US20030165933A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: TANG, Y. Tom

APPLICANT: AZIMZAI, Yalda

APPLICANT: YUE, Henry

APPLICANT: BURFORD, Neil

APPLICANT: DING, Li

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: PATTERSON, Chandra

APPLICANT: BAUGHN, Mariah R.

TITLE OF INVENTION: REGULATORS OF APOPTOSIS

FILE REFERENCE: PI-0307 PCT

CURRENT APPLICATION NUMBER: US/10/296,539

CURRENT FILING DATE: 2002-11-21

PRIOR APPLICATION NUMBER: 60/209,407; 60/250,326

PRIOR FILING DATE: 2000-06-01; 2000-11-30

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PERL Program

SEQ ID NO 4

LENGTH: 1230

TYPE: DNA

ORGANISM: Homo sapiens

FEATURES:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20030165933A1 3102521CB1

US-10-296-539-4

Query Match 99.8%; Score 907.4; DB 16; Length 1230;

Best Local Similarity 99.9%; Pred. No. 5.9e-228;

Matches 908; Conservative 0; Mismatches 1; Indels 0;

QY 1 ATGGCGCTATCCGAGTGCACCCCGGCGCTGCTGGAGAGATGATGCTTGGACTAC 60
 DB 127 ATGGCGCTATCCGAGTGCACCCCGGCGCTGCTGGAGAGATGATGCTTGGACTAC 186
 QY 61 TAGGGAGTGTGTGCTTCAACCGATGTTGAGAGTGGTGGGCGGCAACTGACCGAGTGC 120
 DB 187 TAGGGAGTGTGTGCTTCAACCGATGTTGAGAGTGGTGGGCGGCAACTGACCGAGTGC 246
 QY 121 GAGCTGAGAGCTCTGAGCTTCTGCTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCC 180
 DB 247 GAGCTGAGAGCTCTGAGCTTCTGCTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCC 306
 QY 181 GGGGCGCGCAGCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGTGGCGGAG 240
 DB 307 GGGGCGCGCAGCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGTGGCGGAG 366
 QY 241 AGCAACTGCGGGCTGCTGGGGCAACTCTGCGGGTGTGAGCGCGCGCAAGCTGAGCGG 300
 DB 367 AGCAACTGCGGGCTGCTGGGGCAACTCTGCGGGTGTGAGCGCGCGCAAGCTGAGCGG 426
 QY 301 CACCTGGCGCGCAAGCGCGCGCGCGCAGTGTCTCAGAAAGCTATAGCTATGGCACTCC 360
 DB 427 CACCTGGCGCGCAAGCGCGCGCGCGCAGTGTCTCAGAAAGCTATAGCTATGGCACTCC 486
 QY 361 AGCTTTCAAAAGAGACAGAGGGTGTGCTGCGCGCTGCGGCAATCAAGCAAGTTTGC 420
 DB 487 AGCTTTCAAAAGAGACAGAGGGTGTGCTGCGCGCTGCGGCAATCAAGCAAGTTTGC 546
 QY 421 AATTCTCAGAGGGTTCAGTGGAGACAGAGGCTCCCGCCCAACCAAGGCGGCGGAGT 480
 DB 547 AATTCTCAGAGGGTTCAGTGGAGACAGAGGCTCCCGCCCAACCAAGGCGGCGGAGT 606
 QY 481 CGGGCGCGCGCCAGTGTGTGTCAGACCGCGCGGAGAGAGGGCGCCAGCGGACCCAG 540
 DB 607 CGGGCGCGCGCCAGTGTGTGTCAGACCGCGCGGAGAGAGGGCGCCAGCGGACCCAG 666
 QY 541 CAGCAGTCAGAGCGCGCCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCGCTC 600
 DB 667 CAGCAGTCAGAGCGCGCCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCGCTC 726
 QY 601 CGGGTTGAGACAGAGTCTGCGAGCATGGGCGAGCTTGGAGAGCGGCGTGGCATCCCG 660
 DB 727 CGGGTTGAGACAGAGTCTGCGAGCATGGGCGAGCTTGGAGAGCGGCGTGGCATCCCG 786
 QY 661 CGGCGCGCGAGCGTGGCGCGCAGCTGAGCATGTGTTGGGAGGCGCACCGAGTGTGCGC 720
 DB 787 CGGCGCGCGAGCGTGGCGCGCAGCTGAGCATGTGTTGGGAGGCGCACCGAGTGTGCGC 846
 QY 721 TCAGAGGAGCTGGGCTGTGTGTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 780
 DB 847 TCAGAGGAGCTGGGCTGTGTGTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 906
 QY 781 GCGCTTGGGGCGACATCACTGAGTGGGCGCTGCTGAGGCGCGGCGGGGCGTGTCTCG 840
 DB 907 GCGCTTGGGGCGACATCACTGAGTGGGCGCTGCTGAGGCGCGGCGGGGCGTGTCTCG 966
 QY 841 ACTGAGGCGCTGCGAGAGAGCTGTGTGGCGCGGAGGCTTTCGCTGTGATCAAGTGGAT 900
 DB 967 ACTGAGGCGCTGCGAGAGAGCTGTGTGGCGCGGAGGCTTTCGCTGTGATCAAGTGGAT 1026
 QY 901 GAGGCTGAC 909
 DB 1027 GAGGCTGAC 1035

RESULT 2

US-10-001-254-17
 ; Sequence 17, Application US/10001254
 ; Publication No. US20030049702A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; APPLICANT: Godzik, Adam

; APPLICANT: Pawlowski, Krzysztof
 ; APPLICANT: Fiorentino, Loredana
 ; APPLICANT: Lee, Sung Hyung
 ; APPLICANT: Roth, Wilfried
 ; APPLICANT: Steiner-Liewen, Frank
 ; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
 ; FILE REFERENCE: P-LJ 5037
 ; CURRENT APPLICATION NUMBER: US/10/001,254
 ; CURRENT FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: 60/301,889
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/715,893
 ; PRIOR FILING DATE: 2000-11-17
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 1924
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (91)...(1044)
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(1900)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-001-254-17

Query Match 99.8%; Score 907.4; DB 14; Length 1924;
 Best Local Similarity 99.9%; Pred. No. 5,9e-228;
 Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGAGTGCACCCCGGCGCTGCTGGAGAGATGATGCTTGGACTAC 60
 DB 91 ATGGCGCTATCCGAGTGCACCCCGGCGCTGCTGGAGAGATGATGCTTGGACTAC 150
 QY 61 TAGGGAGTGTGTGCTTCAACCGATGTTGAGAGTGGTGGGCGGCAACTGACCGAGTGC 120
 DB 151 TAGGGAGTGTGTGCTTCAACCGATGTTGAGAGTGGTGGGCGGCAACTGACCGAGTGC 210
 QY 121 GAGCTGAGAGCTCTGAGCTTCTGCTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCC 180
 DB 211 GAGCTGAGAGCTCTGAGCTTCTGCTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCC 270
 QY 181 GGGGCGCGCAGCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGTGGCGGAG 240
 DB 271 GGGGCGCGCAGCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGTGGCGGAG 330
 QY 241 AGCAACTGCGGGCTGCTGGGGCAACTCTGCGGGTGTGAGCGCGCGCAAGCTGAGCGG 300
 DB 331 AGCAACTGCGGGCTGCTGGGGCAACTCTGCGGGTGTGAGCGCGCGCAAGCTGAGCGG 390
 QY 301 CACCTGGCGCGCAAGCGCGCGCGCGCAGTGTCTCAGAAAGCTATAGCTATGGCACTCC 360
 DB 391 CACCTGGCGCGCAAGCGCGCGCGCGCAGTGTCTCAGAAAGCTATAGCTATGGCACTCC 450
 QY 361 AGCTTTCAAAAGAGACAGAGGGTGTGCTGCGCGCTGCGGCAATCAAGCAAGTTTGC 420
 DB 451 AGCTTTCAAAAGAGACAGAGGGTGTGCTGCGCGCTGCGGCAATCAAGCAAGTTTGC 510
 QY 421 AATTCTCAGAGGGTTCAGTGGAGACAGAGGCTCCCGCCCAACCAAGGCGGCGGAGT 480
 DB 511 AATTCTCAGAGGGTTCAGTGGAGACAGAGGCTCCCGCCCAACCAAGGCGGCGGAGT 570
 QY 481 CGGGCGCGCGCCAGTGTGTGTCAGACCGCGCGGAGAGAGGGCGCCAGCGGACCCAG 540
 DB 571 CGGGCGCGCGCCAGTGTGTGTCAGACCGCGCGGAGAGAGGGCGCCAGCGGACCCAG 630
 QY 541 CAGCAGTCAGAGCGCGCCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCGCTC 600
 DB 631 CAGCAGTCAGAGCGCGCCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCGCTC 690
 QY 601 CGGGTTGAGACAGAGTCTGCGAGCATGGGCGAGCTTGGAGAGCGGCGTGGCATCCCG 660

Dh 691 CGGGTTGAGACAGTACTGCGAGCATGAGCCAGCCTTGAGAGCAGGGCGTGGCATCCCGG 750
Qy 661 CGGCCCCAGGCGCTGGGCGGCGAGCTGGACGTGTTTGGGCGAGGCCACGCGAGTCTGCGC 720
Db 751 CGGCCCCAGGCGCTGGGCGGCGAGCTGGACGTGTTTGGGCGAGGCCACGCGAGTCTGCGC 810
Qy 721 TCAAGGAGCTGGGCTCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAC 780
Db 811 TCAAGGAGCTGGGCTCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAC 870
Qy 781 GCCTTCGGGGCGCATCTGAGTGGGCGCCCTGCTGACAGGCCCTGGCGGGGCGTTCCTG 840
Db 871 GCCTTCGGGGCGCATCTGAGTGGGCGCCCTGCTGACAGGCCCTGGCGGGGCGTTCCTG 930
Qy 841 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTGCTGCTGTGATGTGAT 900
Db 931 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTGCTGCTGTGATGTGAT 990
Qy 901 GAGGCTGAC 909
Db 991 GAGGCTGAC 999

RESULT 3

US-10-013-477-2
; Sequence 2, Application US/10013477
; Publication No. US20030049732A1
; GENERAL INFORMATION:
; APPLICANT: Nt et al.
; TITLE OF INVENTION: Apoptosis Related Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO02P1
; CURRENT APPLICATION NUMBER: US/10/013,477
; PRIOR FILING DATE: 2001-12-13
; PRIOR APPLICATION NUMBER: 09/669,445
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: PCT/US00/06642
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/126,018
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/139,638
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/149,449
; PRIOR FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-013-477-2

Query Match 99.8%; Score 907.4; DB 14; Length 2045;

Best Local Similarity 99.9%; Pred. No. 5.9e-228; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCTATCCGGGTCGACCCCGGCGCGTGTGAGAGAGATGAGTGGCTGACTAC 60
Db 121 ATGGCGCTATCCGGGTCGACCCCGGCGCGTGTGAGAGAGATGAGTGGCTGACTAC 180
Qy 61 TACGGAGTGTCTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
Db 181 TACGGAGTGTCTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 240
Qy 121 GAGCTGAGTCTCTGGGCTTTTGTGCTGAGTGAAGCTCTTGGCGCGCGGAGCTTAGCC 180
Db 241 GAGCTGAGTCTCTGGGCTTTTGTGCTGAGTGAAGCTCTTGGCGCGCGGAGCTTAGCC 300
Qy 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCGCGCGGAGAGTGGCGGAG 240
Db 301 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCGCGCGGAGAGTGGCGGAG 360
Qy 241 AGCAACCTGGCGCTGCTGGGCAACTCTCTGCGCTGTGGCGCGCAGACTGCTGCGG 300

Dh 361 AGCAACCTGGCGCTGCTGGGCAACTCTCTGCGCTGTGGCGCGCAGACTGCTGCGG 420
Qy 301 CACCTGCGCGCAAGCGGCGCGGCGAGTGTCTTCAGAACGCTATAGGACCTTC 360
Db 421 CACCTGCGCGCAAGCGGCGCGGCGAGTGTCTTCAGAACGCTATAGGACCTTC 480
Qy 361 AGCTTTCAAGAGAGCAGAGGATAGCTGCGCTGCGCGTGGCAGTCAAGCATTTCTGCA 420
Db 481 AGCTTTCAAGAGAGCAGAGGATAGCTGCGCTGCGCGTGGCAGTCAAGCATTTCTGCA 540
Qy 421 AATTCTCAGAGAGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGCGCGGAGT 480
Db 541 AATTCTCAGAGAGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGCGGAGT 600
Qy 481 CGGGGCGGCGCAGTGTGTGGTGCAGACGCGCGGAGAGAGGAGGCGCCAGCGCACCCAG 540
Db 601 CGGGGCGGCGCAGTGTGTGGTGCAGACGCGCGGAGAGAGGAGGCGCCAGCGCACCCAG 660
Qy 601 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGACCTGTGACATCCGCTC 600
Db 661 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGACCTGTGACATCCGCTC 720
Qy 601 CGGCTTGAAGAGATCTGCGAGCATGGGCGAGCTTGGAGCAGGGCGTGGCATCCCGG 660
Db 721 CGGCTTGAAGAGATCTGCGAGCATGGGCGAGCTTGGAGCAGGGCGTGGCATCCCGG 780
Qy 661 CGGCCCCAGGCGCTGGGCGGCGAGCTGGAAGTGTGGGAGGCGCACCGAGTCTGCGC 720
Db 781 CGGCCCCAGGCGCTGGGCGGCGAGCTGGAAGTGTGGGAGGCGCACCGAGTCTGCGC 840
Qy 721 TCAAGGAGCTGGGCTCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAC 780
Db 841 TCAAGGAGCTGGGCTCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAC 900
Qy 781 GCCTTCGGGGCGCATCTGAGTGGGCGCCCTGCTGACAGGCCCTGGCGGGCGTTCCTG 840
Db 901 GCCTTCGGGGCGCATCTGAGTGGGCGCCCTGCTGACAGGCCCTGGCGGGCGTTCCTG 960
Qy 841 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTGCGCTGTGATGTGAT 900
Db 961 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTGCGCTGTGATGTGAT 1020
Qy 901 GAGGCTGAC 909
Db 1021 GAGGCTGAC 1029

RESULT 4

US-10-106-698-349
; Sequence 349, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 349
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-349

Query Match 99.8%; Score 907.4; DB 15; Length 2045;
Best Local Similarity 99.9%; Pred. No. 5.9e-228;

```
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGTGGAGAGATAGTCTTGACTAC 60
Db 121 ATGGCGCTATCCGGGTGACCCCGCCGCTGTGGAGAGATAGTCTTGACTAC 180
QY 61 TACGGGATGCTGTGCTTCAACCGATGTTGAGAGTGTGGGCACTGACCGAGTGC 120
Db 181 TACGGGATGCTGTGCTTCAACCGATGTTGAGAGTGTGGGCACTGACCGAGTGC 240
QY 121 GAGCTGAGCTCTGTGCTTTCTGTGATGAGAGCTCTTGCCGCGCGAGCTTAGCC 180
Db 241 GAGCTGAGCTCTGTGCTTTCTGTGATGAGAGCTCTTGCCGCGCGAGCTTAGCC 300
QY 181 CGGCGCCGCAAGCGGCTTGAAGCTCTGTGAGAGTGAAGCGCGCGAGCGCGAG 240
Db 301 CGGCGCCGCAAGCGGCTTGAAGCTCTGTGAGAGTGAAGCGCGCGAGCGCGAG 360
QY 241 AGCAACCTGCGGCTGTGGGGCAACTCTGCGGCTGTGCGCCGCAAGCTGTGCGG 300
Db 361 AGCAACCTGCGGCTGTGGGGCAACTCTGCGGCTGTGCGCCGCAAGCTGTGCGG 420
QY 301 CACCTGGCGCGCAAGCGGCGCGCCGCAAGTGTCTCCAGAAAGCTATAGCTACCTCC 360
Db 421 CACCTGGCGCGCAAGCGGCGCGCCGCAAGTGTCTCCAGAAAGCTATAGCTACCTCC 480
QY 361 AGCTTTCAAAGAGACAGAGGGTAGCTGCGCGCTGTGCGAGTCAAGCAATTTGCA 420
Db 481 AGCTTTCAAAGAGACAGAGGGTAGCTGCGCGCTGTGCGAGTCAAGCAATTTGCA 540
QY 421 AATTCTCAGAGGGTAGTGGGAGACAGAGCTCCCCCAACCAAGGGGCGCGCGAGT 480
Db 541 AATTCTCAGAGGGTAGTGGGAGACAGAGCTCCCCCAACCAAGGGGCGCGCGAGT 600
QY 481 CGGCGCCGCGCAAGTGTGTGCGCAGACGCGCGGAGAGGGGCGCCAGCGCACCCAG 540
Db 601 CGGCGCCGCGCAAGTGTGTGCGCAGACGCGCGGAGAGGGGCGCCAGCGCACCCAG 660
QY 601 CAGCACTCAGAGCGCCCGCAGACCTTCTTGAAGGCAAGTGAATCTGATCCGCTC 720
Db 661 CAGCACTCAGAGCGCCCGCAGACCTTCTTGAAGGCAAGTGAATCTGATCCGCTC 780
QY 661 CGGCTTCAGAGAGTACGTGCGAGCATGGGCGAGCTGTGAGAGAGGCGTGGATCCCG 600
Db 721 CGGCTTCAGAGAGTACGTGCGAGCATGGGCGAGCTGTGAGAGAGGCGTGGATCCCG 660
QY 721 TCAGAGGACCTGGGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTCTATCTGAG 780
Db 841 TCAGAGGACCTGGGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTCTATCTGAG 900
QY 841 ACTGAGGACCTGGGCTGTGAGAGCTGTGGGCGGAGAGGCTGTGCTGTGATGAT 900
Db 961 ACTGAGGACCTGGGCTGTGAGAGCTGTGGGCGGAGAGGCTGTGCTGTGATGAT 1020
QY 901 GAGGCTGAG 909
Db 1021 GAGGCTGAG 1029
```

```

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 315
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-315

Query Match          98.5%; Score 895.4; DB 9; Length 2044;
Best Local Similarity 99.8%; Pred. No. 8,1e-225;
Matches 907; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGTGGAGAGATAGTCTTGACTAC 60
Db 121 ATGGCGCTATCCGGGTGACCCCGCCGCTGTGGAGAGATAGTCTTGACTAC 180
QY 61 TACGGGATGCTGTGCTTCAACCGATGTTGAGAGTGTGGGCACTGACCGAGTGC 120
Db 181 TACGGGATGCTGTGCTTCAACCGATGTTGAGAGTGTGGGCACTGACCGAGTGC 240
QY 121 GAGCTGAGCTCTGTGCTTTCTGTGATGAGAGCTCTTGCCGCGCGAGCTTAGCC 180
Db 241 GAGCTGAGCTCTGTGCTTTCTGTGATGAGAGCTCTTGCCGCGCGAGCTTAGCC 300
QY 181 CGGCGCCGCAAGCGGCTTGAAGCTCTGTGAGAGTGAAGCGCGCGAGCGCGAG 240
Db 301 CGGCGCCGCAAGCGGCTTGAAGCTCTGTGAGAGTGAAGCGCGCGAGCGCGAG 360
QY 241 AGCAACCTGCGGCTGTGGGGCAACTCTGCGGCTGTGCGCCGCAAGCTGTGCGG 300
Db 361 AGCAACCTGCGGCTGTGGGGCAACTCTGCGGCTGTGCGCCGCAAGCTGTGCGG 420
QY 301 CACCTGGCGCGCAAGCGGCGCGCCGCAAGTGTCTCCAGAAAGCTATAGCTACCTCC 360
Db 421 CACCTGGCGCGCAAGCGGCGCGCCGCAAGTGTCTCCAGAAAGCTATAGCTACCTCC 480
QY 361 AGCTTTCAAAGAGACAGAGGGTAGCTGCGCGCTGTGCGAGTCAAGCAATTTGCA 420
Db 481 AGCTTTCAAAGAGACAGAGGGTAGCTGCGCGCTGTGCGAGTCAAGCAATTTGCA 540
QY 421 AATTCTCAGAGGGTAGTGGGAGACAGAGCTCCCCCAACCAAGGGGCGCGCGAGT 480
Db 541 AATTCTCAGAGGGTAGTGGGAGACAGAGCTCCCCCAACCAAGGGGCGCGCGAGT 600
QY 541 CGGCGCCGCGCAAGTGTGTGCGCAGACGCGCGGAGAGGGGCGCCAGCGCACCCAG 540
Db 601 CGGCGCCGCGCAAGTGTGTGCGCAGACGCGCGGAGAGGGGCGCCAGCGCACCCAG 660
QY 601 CAGCACTCAGAGCGCCCGCAGACCTTCTTGAAGGCAAGTGAATCTGATCCGCTC 720
Db 661 CAGCACTCAGAGCGCCCGCAGACCTTCTTGAAGGCAAGTGAATCTGATCCGCTC 780
QY 661 CGGCTTCAGAGAGTACGTGCGAGCATGGGCGAGCTGTGAGAGAGGCGTGGATCCCG 600
Db 721 CGGCTTCAGAGAGTACGTGCGAGCATGGGCGAGCTGTGAGAGAGGCGTGGATCCCG 660
QY 721 TCAGAGGACCTGGGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTCTATCTGAG 780
Db 841 TCAGAGGACCTGGGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTCTATCTGAG 900
QY 841 ACTGAGGACCTGGGCTGTGAGAGCTGTGGGCGGAGAGGCTGTGCTGTGATGAT 900
Db 961 ACTGAGGACCTGGGCTGTGAGAGCTGTGGGCGGAGAGGCTGTGCTGTGATGAT 1020
QY 901 GAGGCTGAG 909
Db 1021 GAGGCTGAG 1029
```

```
RESULT 5
US-09-925-302-315
; Sequence 315, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
```


Db 901 GCCTTCGGGGCGACTGAGTGGCGCCCTGCTGCA-GCCCTGCGGGGGTGTCTCG 959
Qy 841 ACTGAGCCCTGCGAGAGGCTGTGGCGGGAGGCTGTGGCTGTGTCAGTGGAT 900
Db 960 ACTGAGCCCTGCGAGAGGCTGTGGCGGGAGGCTGTGGCTGTGTCAGTGGAT 1019
Qy 901 GAGGCTGAC 909
Db 1020 GAGGCTGAC 1028

RESULT 6
US-09-925-302-315

; Sequence 315, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 315
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-315

Query Match 98.5%; Score 895.4; DB 10; Length 2044;

Best Local Similarity 99.8%; Pred. No. 8,1e-225;
Matches 907; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGGAGAGATGAGTGCCTGACTAC 60
Db 121 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGGAGAGATGAGTGCCTGACTAC 180
Qy 61 TAGCGGATGCTGTGCTTCAACCGTATGTCGAGTGGCGGGCAACTGACCGAGTGC 120
Db 181 TAGCGGATGCTGTGCTTCAACCGTATGTCGAGTGGCGGGCAACTGACCGAGTGC 240
Qy 121 GAGCTGAGCTCTCTGCTTCTGCTGATGAGGCTCTTGCGCGCCGAGAGCTTACC 180
Db 241 GAGCTGAGCTCTCTGCTTCTGCTGATGAGGCTCTTGCGCGCCGAGAGCTTACC 300
Qy 181 CGGGCCGCGAGCGGCTAGAGCTCTGCTGAGTGGAGCGCGGGCAAGTGGCGGAG 240
Db 301 CGGGCCGCGAGCGGCTAGAGCTCTGCTGAGTGGAGCGCGGGCAAGTGGCGGAG 360
Qy 241 AGCAACCTGGCGCTGCTGGGCACTCTGCGCTGCTGGCGCCGCAAGCTTGCTCG 300
Db 361 AGCAACCTGGCGCTGCTGGGCACTCTGCGCTGCTGGCGCCGCAAGCTTGCTCG 420
Qy 301 CACTGCGCGCAAGCGCGCGCGCAAGTGTCTCAGAACTTACTATGAGCACTTCC 480
Db 421 CACTGCGCGCAAGCGCGCGCGCAAGTGTCTCAGAACTTACTATGAGCACTTCC 480
Qy 361 AGCTCTTCAAGAGAGAGAGTGTGCTGCGCTGCGCGGAGTCAAGAGATTCTGCA 420
Db 481 AGCTCTTCAAGAGAGAGAGTGTGCTGCGCTGCGCGGAGTCAAGAGATTCTGCA 540
Qy 421 AATTCTCAGAGGGTCAAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
Db 541 AATTCTCAGAGGGTCAAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 600
Qy 481 CGGGGCGGGCCAGTGTGTGCTGCAAGCGGCGGAGAGGGGCGCCAGCGCACCCAG 540
Db 601 CGGGGCGGGCCAGTGTGTGCTGCAAGCGGCGGAGAGGGGCGCCAGCGCACCCAG 660

Qy 541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAGTGAATTGATCCGGCTC 600
Db 661 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAGTGAATTGATCCGGCTC 720
Qy 601 CGGGTTGAGACAGATCTGCGAGCATGGGCGAGCTTGGAGAGAGGCGGCAATCCGG 660
Db 721 CGGGTTGAGACAGATCTGCGAGCATGGGCGAGCTTGGAGAGAGGCGGCAATCCGG 780
Qy 661 CGGGCCGAGGCGCTGGCGCGGAGCTGAGAGTGTGGAGAGGCAACGAGTGCCTGC 720
Db 781 CGGGCCGAGGCGCTGGCGCGGAGCTGAGAGTGTGGAGAGGCAACGAGTGCCTGC 840
Qy 721 TCAAGGACCTGGGCTGTGTTGTGATCAAGTTCAGAGCTCTCTATCTGAG 780
Db 841 TCAAGGACCTGGGCTGTGTTGTGATCAAGTTCAGAGCTCTCTATCTGAG 900
Qy 781 GCCTTCGGGGCGACTACTGAGTGGCGCCCTGCTGAGGCTTGGCGGGGCTTCTTG 840
Db 901 GCCTTCGGGGCGACTACTGAGTGGCGCCCTGCTGCA-GCCCTGCGGGCGTGTCTTG 959
Qy 841 ACTGAGCCCTGCGAGAGGCTGTGGCGGGAGGCTGTGGCTGTGTCAGTGGAT 900
Db 960 ACTGAGCCCTGCGAGAGGCTGTGGCGGGAGGCTGTGCTGTGTCAGTGGAT 1019
Qy 901 GAGGCTGAC 909
Db 1020 GAGGCTGAC 1028

RESULT 7
US-09-822-830A-390

; Sequence 390, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fachel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakkar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 390
; LENGTH: 1966
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-830A-390

Query Match 87.3%; Score 799.4; DB 9; Length 1966;
Best Local Similarity 90.2%; Pred. No. 1,1e-199;
Matches 908; Conservative 0; Mismatches 1; Indels 98; Gaps 1;

Qy 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGGAGAGATGAGTGCCTGAGTAC 60
Db 50 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGGAGAGATGAGTGCCTGAGTAC 109
Qy 61 TAGCGGATGCTGTGCTTCAACCGTATGTCGAGTGGTGGCGGCAACTGACGAGTGC 120
Db 110 TAGCGGATGCTGTGCTTCAACCGTATGTCGAGTGGTGGCGGCAACTGACGAGTGC 169
Qy 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTACC 180
Db 170 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTACC 229

```
QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGTGAAGCGCCGCGGCACTGCGCGAG 240
DB 230 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGTGAAGCGCCGCGGCACTGCGCGAG 289
QY 241 AGCAACCTGGGCTGCTGGGGCAACTCTGCGGCTGTGGCCCGCCAGCACTGCTGGCG 300
DB 290 AGCAACCTGGGCTGCTGGGGCAACTCTGCGGCTGTGGCCCGCCAGCACTGCTGGCG 349
QY 301 CACTGCGCGCAGCGGCGCCGCGCAGTGTCTCCAGAACGCTATAGCTATGCGACCTCC 360
DB 350 CACTGCGCGCAGCGGCGCCGCGCAGTGTCTCCAGAACGCTATAGCTATGCGACCTCC 409
QY 361 AGCTCTTCAAAAGAGACAGAGGGTAGCTGCGCTGCGGCTGCGGCACTGCTGCA 420
DB 410 AGCTCTTCAAAAGAGACAGAGGGTAGCTGCGCTGCGGCTGCGGCACTGCTGCA 469
QY 421 AATTCTCAGAGGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGCGCGGAGT 480
DB 470 AATTCTCAGAGGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGCGGAGT 529
QY 481 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGGCGAGAGGGGCCCGCAGCCGACCCAG 540
DB 530 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGGCGAGAGGGGCCCGCAGCCGACCCAG 589
QY 541 CAGCACTCAGAGCCCGCAGACCTTCTCTGAAAGCAAAGTGACTG----- 587
DB 590 CAGCACTCAGAGCCCGCAGACCTTCTCTGAAAGCAAAGTGACTGCTGTACAGA 649
QY 588 ----- 587
DB 650 AGCAGGGTCCAGCATCTCTCTGTTAGAGACCTCCGGAAGCTTCCATTCACTGTGAA 709
QY 588 -----TGACATCCGCGCTCCGCGTTCAGAGCAGAGTACTGCG 622
DB 710 GGGCAAGGGAGAGGCTTGTCAATGACATCCGGCTCCGGGTTCCAGAGCAGAGTACGCG 769
QY 623 AGCATGGGCGACCTTGGAGAGGCGGTGCAATCCGCGCGCCCGCAGGCGCTGCGCGCG 662
DB 770 AGCATGGGCGACCTTGGAGAGGCGGTGCAATCCGCGCGCCCGCAGGCGCTGCGCGCG 829
QY 683 AGCTGACGCTTTGGGCAAGCCACGCGAGTGTGCGCTCAAGGGAACCTGCGGCTGTG 742
DB 830 AGCTGACGCTTTGGGCAAGCCACGCGAGTGTGCGCTCAAGGGAACCTGCGGCTGTG 889
QY 743 TTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGCGGCGCACTACTGA 802
DB 890 TTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGCGGCGCACTACTGA 949
QY 803 GTGGCGCGCTGCTGACAGGCGCTGCGGGCGGTGTTCTGACTGAGGCCCTGCGAGAGGCTG 862
DB 950 GTGGCGCGCTGCTGACAGGCGCTGCGGGCGGTGTTCTGACTGAGGCCCTGCGAGAGGCTG 1009
QY 863 TGGGCGGGGAGGCTTTCGCTGCTGTGCTGAGTGAGTGAGGCTGAC 909
DB 1010 TGGGCGGGGAGGCTTTCGCTGCTGTGCTGAGTGAGTGAGGCTGAC 1056

RESULT 8
US-10-296-115-481
; Sequence 481, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
```

```
; SEQ ID NO 481
; LENGTH: 1570
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-481

Query Match      63.9%; Score 581.2; DB 17; Length 1570;
Best Local Similarity 96.3%; Pred. No. 1,6e-142;
Matches 595; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 292 CTGCTCCGCGACCTGCGCGCAGCGCGCCGCGCAGTGTCTCCAGAACGCTATAGCTAT 351
DB 40 CGGCGCTCCGCGCAGAGTATCGAGAGCCAGCAATCTGTCTCTCAGAACGCTATAGCTAT 99
QY 352 GGCACCTCCAGCTCTTCAAAAGAGACAGAGGGTAGCTGCGCGCTGCGGCACTGCAAGC 411
DB 100 GGCACCTCCAGCTCTTCAAAAGAGACAGAGGGTAGCTGCGCGCTGCGGCACTGCAAGC 159
QY 412 AGTTCTGCAAAATTCAGAGGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGCGAG 471
DB 160 AGTTCTGCAAAATTCAGAGGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGCGAG 219
QY 472 CGGCGGAGTCTGGGCGCGCCAGTGTGTGTCAGACGCGCGGCGAGAGGGGCCCGCAGCC 531
DB 220 CGGCGGAGTCTGGGCGCGCCAGTGTGTGTCAGACGCGCGGCGAGAGGGGCCCGCAGCC 279
QY 532 GCACCCGACAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAAGTGACTGTGAC 591
DB 280 GCACCCGACAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAAGTGACTGTGAC 339
QY 592 ATCCGGCTCCGGTTTGAAGCAGAGTACTGCGAGCAATGGGCGAGGCTTGGAGCAGGGCGTG 651
DB 340 ATCCGGCTCCGGTTTGAAGCAGAGTACTGCGAGCAATGGGCGAGGCTTGGAGCAGGGCGTG 399
QY 652 GCATCCCGGCGGCGCCAGGCGCTGTGCGCGGCACTGGAAGTGTGTTGGAGGCAACCGGA 711
DB 400 GCATCCCGGCGGCGCCAGGCGCTGTGCGCGGCACTGGAAGTGTGTTGGAGGCAACCGGA 459
QY 712 GTGCTGCGCTCAAGGAGACCTGCGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCC 771
DB 460 GTGCTGCGCTCAAGGAGACCTGCGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCC 519
QY 772 TATCTGAGAGCGCTTCTGGGCGCACTACCTGAGTGGGCGCTGCTGAGAGCGCTGCGGGCG 831
DB 520 TATCTGAGAGCGCTTCTGGGCGCACTACCTGAGTGGGCGCTGCTGAGAGCGCTGCGGGCG 579
QY 832 GTGTTCTGACAGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTTTCGCTGCTGTGTC 891
DB 580 GTGTTCTGACAGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTTTCGCTGCTGTGTC 639
QY 892 AGTGTGATGAGGCTGAC 909
DB 640 AGTGTGATGAGGCTGAC 657

RESULT 9
US-10-037-270-853
; Sequence 853, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aildong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, ping
; APPLICANT: Ma, Yundong
```

```

; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 853
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(984)
US-10-037-270-853
```

```

Query Match      59.5%; Score 541; DB 15; Length 1067;
Best Local Similarity 95.8%; Pred. No. 5.3e-132;
Matches 572; Conservative 0; Mismatches 10; Indels 15; Gaps 1;
```

```

QY 1 ATGGCGCTATCCGGGTGACCCCGGCGCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
DB 151 ATGGCGCTATCCGGGTGACCCCGGCGCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 210
QY 61 TACGGAGTGTGCTGCTGCTTCAACCGATGTTGAGAGTGTGGGCGGCACTGACCGAGTGC 120
DB 211 TACGGAGTGTGCTGCTGCTTCAACCGATGTTGAGAGTGTGGGCGGCACTGACCGAGTGC 270
QY 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTAAACC 180
DB 271 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTAAACC 330
QY 181 CGGGCGCGGCGGCTTGAAGCTCTGCTGAGAGCTGAGGCGCGCGGCACTGAGCGAG 240
DB 331 CGGGCGCGGCGGCTTGAAGCTCTGCTGAGAGCTGAGGCGCGCGGCACTGAGCGAG 390
QY 241 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCGGCGGCACTGAGCGAG 300
DB 391 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCGGCGGCACTGAGCGAG 450
QY 301 CACCTGGCGGCGGCACTGCGGCGGCGGCACTGCTGCGGCGGCGGCACTGAGCGAG 360
DB 451 CACCTGGCGGCGGCACTGCGGCGGCGGCGGCACTGCTGCGGCGGCGGCACTGAGCGAG 510
QY 361 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 511 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
QY 421 AATCTCAGCAGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 571 AATCTCAGCAGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
QY 481 CGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
DB 616 CGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 675
QY 541 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
DB 676 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
```

```

RESULT 10
US-10-117-722-853
; Sequence 853; Application US/10117722
; Publication No. US2003021974A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US2003021974A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BICP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 853
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(984)
US-10-117-722-853
```

```

Query Match      59.5%; Score 541; DB 17; Length 1067;
Best Local Similarity 95.8%; Pred. No. 5.3e-132;
Matches 572; Conservative 0; Mismatches 10; Indels 15; Gaps 1;
```

```

QY 1 ATGGCGCTATCCGGGTGACCCCGGCGCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
DB 151 ATGGCGCTATCCGGGTGACCCCGGCGCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 210
QY 61 TACGGAGTGTGCTGCTGCTTCAACCGATGTTGAGAGTGTGGGCGGCACTGACCGAGTGC 120
DB 211 TACGGAGTGTGCTGCTGCTTCAACCGATGTTGAGAGTGTGGGCGGCACTGACCGAGTGC 270
QY 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAAACC 180
DB 271 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAAACC 330
QY 181 CGGGCGCGGCGGCTTGAAGCTCTGCTGAGAGCTGAGGCGCGCGGCACTGAGCGAG 240
DB 331 CGGGCGCGGCGGCTTGAAGCTCTGCTGAGAGCTGAGGCGCGCGGCACTGAGCGAG 390
QY 241 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCGGCGGCACTGAGCGAG 300
DB 391 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCGGCGGCACTGAGCGAG 450
QY 301 CACCTGGCGGCGGCACTGCGGCGGCGGCGGCGGCACTGAGCGAG 360
DB 451 CACCTGGCGGCGGCACTGCGGCGGCGGCGGCGGCACTGAGCGAG 510
QY 361 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 511 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
QY 421 AATCTCAGCAGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 571 AATCTCAGCAGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
QY 481 CGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
DB 616 CGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 675
QY 541 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
DB 676 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
```

```
RESULT 11
US-09-799-777-103
; Sequence 103, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Corley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sacher, Susan
; Shah, Puri
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1554 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRANO706
; CLONE: 1638407
; SEQUENCE DESCRIPTION: SEQ ID NO: 103 :
US-09-799-777-103

Query Match          59.3%; Score 539.2; DB 9; Length 1554;
Best Local Similarity 95.0%; Pred. No. 1,6e-131;
Matches 576; Conservative 0; Mismatches 8; Indels 22; Gaps 1;
```

```

Db      269 CCTCTGAAAGGCAAGGTGACCTGTGATGGGAAAACAGGCTCAGAGAGACATCCGGCTCCGG 328
Qy      604 GTTCGACGAGATGACTGCGAGATGGGCCAGCCTTGAGACAGGCGCTGGCATCCGCGCG 663
Db      329 GTTCGAGCAGAGTACTGCGAGCATGGGCAGCCTTGAGACAGGCGCTGGCATCCGCGCG 388
Qy      664 CCCGAGCGCTGGCGCGGCGAGCTGAGCGTGTGGGCGAGGCCACCGCATGCTGCGCTCA 723
Db      389 CCCGAGCGCTGGCGCGGCGAGCTGAGCGTGTGGGCGAGGCCACCGCATGCTGCGCTCA 448
Qy      724 AGGGAACCTGAGGCTGTGGTTGTGATCATCAAGTTCTCAGAGCTCCTATCTGACGCC 783
Db      449 AGGGAACCTGAGGCTGTGGTTGTGATCATCAAGTTCTCAGAGCTCCTATCTGACGCC 508
Qy      784 TTCTGGGGGAGCTACCTGAGTGGCGCCCTGTCGACAGCCCTGCGGGCGTGTCTGACT 843
Db      509 TTCTGGGGGAGCTACCTGAGTGGCGCCCTGTCGACAGCCCTGCGGGCGTGTCTGACT 568
Qy      844 GAGGCGCTGCGAGAGGCTGTGGGCCGCGAGGCTGTTCGCTGCTGTGATGATGAG 903
Db      569 GAGGCGCTGCGAGAGGCTGTGGGCCGCGAGGCTGTTCGCTGCTGTGATGATGAG 628
Qy      904 GCTGAC 909
Db      629 GCTGAC 634

RESULT 12
US-10-723-860-1627/c
; Sequence 1627, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05862.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1627
; LENGTH: 111084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1627

Query Match          35.9%; Score 326.6; DB 18; Length 111084;
Best Local Similarity 97.4%; Pred. No. 6.9e-76;
Matches 332; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

Qy	Dy	Dd
301	CACCTGGCGCGCAAGCGGCGCGGTCAGTGTCTCCAGAAG	341
8562	CACCTGGCGCGCAAGCGGCGCGGTCAGTGTCTCCAGAAG	8522

```

RESULT 13
US-10-001-254-7
; Sequence 7, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfred
; APPLICANT: Stenmer-Liessen, Frank
; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
; FILE REFERENCE: P-LU 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(303)
; US-10-001-254-7

```

	Query Match	33.2%;	Score 301.4;	DB 14;	Length 303;	
	Best Local Similarity	99.7%;	Pred. No. 3e-69;			
	Matches 302;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0.	
QY	34 TGGAGGAGAGATGAGTGCCTGSACTACTACCGGAGTGCTGTCCCTTCACCGATATTGCCAG	93				
Dd	1 TGGAGGAGGATGAGTGCCTGSACTACTACCGGAGTGCTGTCCCTTCACCGATATTGCCAG	60				
QY	94 GTGATGGGCGGGCAACTGACCGAATGCGAAGCTTGGAGCTTTCTGCTGGATGAG	153				
Dd	61 GTGATGGGCGGGCAACTGACCGAATGCGAAGCTTGGAGCTTTCTGCTGGATGAG	120				
QY	154 GCTCCTGGCGCGCGCGAGGCTTAAGCCCGGGGCCGAGCGGGCTTAGAACCTCTGTGGAG	213				
Dd	121 GCTCCTGGCGCGCGCGAGGCTTAAGCCCGGGGCCGAGCGGGCTTAGAACCTCTGTGGAG	180				
QY	214 CTGAGAGCGCCGCGGSGCATGTGCGGCGAGAGCAACTGCGGCTGTGGGCAACTCCTGGCG	273				
Dd	181 CTGAGAGCGCGCGGSGCATGTGCGGCGAGAGCAACTGCGGCTGTGTGGGCAACTCCTGGCG	240				
QY	274 GTGCTGGCGCGCCACAGACTGTGTCGCGCACTTGCGCGCAAGCGGCGCGSACAGTGTCT	333				
Dd	241 GTGCTGGCGCGCCACAGACTGTGTCGCGCACTTGCGCGCAAGCGGCGCGSACAGTGTCT	300				
QY	334 CCA 336					
Dd	301 CCA 303					

RESULT 14
US-09-733-167-2
; Sequence 2, Application US/09733167
; Patent No. US20020099009A1
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus

```

1 APPLICANT: Krammer, Peter
2 TITLE OF INVENTION: Protein for Regulation of Apoptosis
3 FILE REFERENCE: 4121-120
4 CURRENT APPLICATION NUMBER: US/09/733,167
5 CURRENT FILING DATE: 2000-12-08
6 PRIOR APPLICATION NUMBER: PCT/DE99/01712
7 PRIOR FILING DATE: 1999-06-08
8 PRIOR APPLICATION NUMBER: German Patent Application No. US2002009909A1 198 25 621.3
9 PRIOR FILING DATE: 1998-06-08
10 NUMBER OF SEQ ID NOS: 8
11 SOFTWARE: PatentIn version 3.1
12 SEQ ID NO 2
13 LENGTH: 1142
14 TYPE: DNA
15 ORGANISM: Homo sapiens
16 FEATURE:
17 NAME/KEY: misc feature
18 LOCATION: (1042)..(1042)
19 OTHER INFORMATION: n represents any one of a, c, t, and g.
20 NAME/KEY: misc_feature
21 LOCATION: (1043)..(1043)
22 OTHER INFORMATION: n represents any one of a, c, t, and g.
23 NAME/KEY: misc_feature
24 LOCATION: (1114)..(1114)
25 OTHER INFORMATION: n represents any one of a, c, t, and g.
26 US-09-733-167-2

```

Query Match	21.3%;	Score 193.8;	DB 9;	Length 1142;
Best Local Similarity	-52.9%;	Pred. No. 4.5e-41;		
Matches 467;	Conservative 0;	Mismatches 407;	Indels 9;	Gaps 2

QY	27	CCCGTCGTGGAGGAGGATGATGCTCCCTGAACTTAACTGAGATGCTGCTCAACCGTAT	86
Db	54	CGAGGTGTGGCAGAAAGACATGGTGTAGCAGGAAACACGGGCTGTACACCTGCACCGCAT	113
QY	87	GTTTCGAGGTGGTGGGCGGGCACTGACCCAGTCCGAGCTGAGACTCTGGGCTTTCTGCT	146
Db	114	GTTTGACATCGTGGGCACTCATCTGAGACACACAGAGATGTGCGGCTTTCTTCTCTCT	173
QY	147	GGATGAGAGCTTCCTGGCGCGCGCGAGGCTTGACCCGGGCGCGCAGCGGCTTAGAGCTCT	206
Db	174	TGTTGATGTCAATTATATGACACACGAGCGTGAATCTAT--CCGAAATGAGACGTGACTTCTT	230
QY	207	GCTGAGACTGAGCGCGCGCGGAGTGCAGGAGCAACTGCGGCTGTGGGGCAACT	266
Db	231	ATTGGCACTGGAGCGCGCAGGGCCCCCTGTGATGAAAGTAATCTTGGCCAGGTGCTCAGCT	290
QY	267	CTGCGGCGTGTGGGCGCGGCAACGACCTGTGCGGCCACCTGGCGGGCAAGCGGCGCGCG	326
Db	291	GCTGCGGATATCACTGCGCACGACTGTGTGCCCTTACGTCACCTCAAGAGGAGCGGGG	350
QY	327	AGTGTCTCAAAACGCTAATAGCTATGGCACCCTCAGACTCTTCAAAAGAGACAGAGGTATG	386
Db	351	TGTGTGCCCTGTGATCTTGTGACAGATATCTGAGAGACATCAATTTGCTATGTGACCCC	410
QY	387	CTGCGCTGTGCGGTGTGGGACATCAAGCAATTTCTGCAATTTCTACGAGGGTCACTGGGAAC	446
Db	411	CAGAGCCTTAGTATCAAGAACCAAGAGCC----TCCCAAGCCCTTAAACAGTGC	464
QY	447	AGGCTCCCCCAACCAAGCGGCGACGCGGAGTCTGGGGCGGAGCCAGTGTGTGTCAG	506
Db	465	TCCCCACTATTCGTGTGTGTGTGTCCCACTTGGGGTCTCAGATGTGTACCAAGCGGCC	524
QY	507	ACGGCGCGGAGAGAGGGGCCCAAGCCGACCCCAAGACATCAAGACCCCGCCACACTTTC	566
Db	525	AGCCCGAGGAGAACCAACTTTGGAGACCAAGCGAAACCGCGGAAGTATGATGACATCCAGA	584
QY	567	CTCTGAAGGCAAAATGATCTGTGACATCCGGGCTCCGGGTTGAGACAGATGACTGAGACA	626
Db	585	TCCCAAGAGAGAGACATGTGACATCAATGACTGGGGTTTGGGGCTGAATACTGCGACGA	644
QY	627	TGGGCCAGCCTTGGAGCAGGGCGGTGCATTCGGGCGGCCCAAGGCGCTGCGCGGACGCT	686

Db 645 TGAAGACTGCTCTGACAGGCAATGTCTTCTTAACAAGACAGACCACTTGAAGGCCAGTT 704
 QY 687 GGACGTGTTTGGGACAGCCACCGCAGTGTGCGCTCAAGGACCTGAGGCTGTGTGTTG 746
 Db 705 TGAAGCCTTTAAACCAAGCCAAACCAATCTCAAGTCCCGGAGCTGGGCTCCATCATCTG 764
 QY 747 TGAATCAAGTTTCTCAGAGCTCTCTTATCTGGAAGCGCTTCTGGGGGAGTACTAGTGTG 806
 Db 765 TGACATCAAGTTCTCGAGCTCACCTACCTCGATGCAATCTGGGCTGACTTACATCATGG 824
 QY 807 CGCCCTGCTGACAGCCCTGCGGGCGGTCTCTGACTGAGAGCCCTGCGAGAGGCTGTGG 866
 Db 825 CTCTTTATTAGAGGCACTTAAGGTGTCTTTCATCAAGAGACTCCCTCAAGAGACTGTGG 884
 QY 867 CCGGAGAGCTGTTCGCTGTGTGTCAGTGTGATGAGGCTGAC 909
 Db 885 CCAATGAAGCCATCAAGCTGTGTGTTAATGTAGACGAGAGGAC 927

RESULT 15
 US-09-920-300A-1684
 ; Sequence 1684, Application US/09920300A
 ; Patent No. US20020136728A1
 ; GENERAL INFORMATION:
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Secrist, Heather
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.547
 ; CURRENT APPLICATION NUMBER: US/09/920.300A
 ; CURRENT FILING DATE: 2001-07-31
 ; NUMBER OF SEQ ID NOS: 1789
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 1684
 ; LENGTH: 2261
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-920-300A-1684

Query Match 21.3%; Score 193.8; DB 9; Length 2261;
 Best Local Similarity 52.9%; Pred. No. 4,4e-41;
 Matches 467; Conservative 0; Mismatches 407; Indels 9; Gaps 2;

QY 27 CCCGTGCTGGAGAGATGATGCTGACTACTACGAGATGCTGCTGACCGCTAT 86
 Db 165 CCAAGTGTGCGCAAGAGCATGTGAGCAGAAACATGGGCTGTACAGCTGACCGCAT 224
 QY 87 GTTCGAGGTGTGGCGGGCACTGACCGAGTGTGAGCTGTGAGCTCTTCTGCT 146
 Db 225 GTTTGACATGTGGGCACTCATCTGACACAGAGATGTGGGCTTCTTCTCTT 284
 QY 147 GGATGAGGCTCTGAGCGCCCGCGGAGCTTAGCCCGGCGCGGAGCTGCT 206
 Db 285 TGTGTGATGTATGTATGATACCAAGAGCTGTGACTC--ATCCGAATGTGACGTTCTT 341
 QY 207 GCTGAGACTGAGAGCGCGCGGAGTGTGCGGAGAGAGCAACTGTGCGGCTGTGAGGCACT 266
 Db 342 ATTGCACTGTGAGGCGCGGCGGCTGTGATGAAGTAACTTTGCGCAGGTGCTGAGCT 401
 QY 267 CTTGCGGCTGTGCGCGGAGCACTGTGCGGACCTGCGGCGGAGCGCGCGGCG 326
 Db 402 GCTGCGCATCATCTGCGGAGCTGTGCGCTTACCTCAAGAGAGAGAGCGGGC 461
 QY 327 AGTGTCTCCAGAGCGTATAGTATGGAAGTGTGCAAGTGTGCAAGAGAGAGGAGT 386
 Db 462 TGTGTGCTGTGATCTTGTAGCAAGTATCTGAGAGAGACATCAATTCGCTATGTGACCC 521
 QY 387 CTGCGCTGTGCGGAGTGTGAGCAAGTGTGCAATTTCTCAGCAGGAGTCAAGTGGAGAC 446
 Db 522 CAGAGCCCTCATGTATCAGAACCAAGGCC-----TCCCGAGCCCTTAAACAGAGTCC 575

QY 447 AGGCTCCCCCAACCAAGCGGAGCGGCGAGTCCGGGCGCGCCCAAGTGTGTGCCAG 506
 Db 576 TCCCACTATCTGTGTGTGTGTGTCCTCCACTTCGGGTCTCCGAATGTGTAGCAAGCGGCC 635
 QY 507 ACGGCGGCGAGAGAGGAGGCCCAAGCGGCAAGCCAGAGCAGTGAAGCCCGGAGACTTTC 566
 Db 636 AGCCGAGGAGAGAGCCACACTTGGGAGCCAGGAAAGCGCGGAGTCAATGACACACAGA 695
 QY 567 CTCTGAAGCAAGTGAAGTGTGACATCCGGCTCCGGGTTTCCAGACAGATACTGCGAGCA 626
 Db 696 TCCCAAGGAGAAAGACAGATGTGACATCAGACTGCGGGGTTCCGGCTGAAATCTGCGAGCA 755
 QY 627 TGGGCGAGCCTTGAAGCAAGGAGTGGCATCCCGGCGGCCCAAGGCTGCGCGGAGCT 686
 Db 756 TGAAGTCTCTGCAAGGSCAATGTCTTCTTAACAAGCAGAACCACTTGAAGCGCAATT 815
 QY 687 GGACGTGTTTGGGACAGGCCACCGCAGTGTGCGCTCAAGGAGACCTGGGCTGTGTGTTG 746
 Db 816 TGAAGCCTTTAAACAGAGCCAAACCATCTCAAGTCCGGGACCTGGGCTCCATCATCTG 875
 QY 747 TGAATCAAGTTCTCAGAGCTCTCTTATCTGGAAGCGCTTCTGGGCGCACTTACTGAGTGG 806
 Db 876 TGACATCAAGTTCTTGAAGCTCACCTCGATGCAATTCGGCGTGAATCATCATGG 935
 QY 807 CGCCCTGCTGACAGGCCCTGCGGGCGGTTCCTGAAGTGAAGGCCCTGCGAGAGGCTGTGG 866
 Db 936 CTCTTTATTAGAGGCACTTAAGGTGTCTTTCATCACAGACTCCCTCAAGCAAGCTGTGG 995
 QY 867 CCGGAGAGCTGTTCGCTGTGTGTCAGTGTGATGAGGCTGAC 909
 Db 996 CCAATGAAGCCATCAAGCTGTGTGTTAATGTAGACGAGAGGAC 1038

Search completed: March 23, 2005, 17:06:26
 Job time : 573.067 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 23, 2005, 05:31:52 ; Search time 526.452 Seconds
(without alignments)
10221.347 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctaccgcggctgcac.....tcagtgctgacgagctgac 909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapect 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001as:*
5: geneseq2001bs:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	909	100.0	909	5	AAF27407 Human apo
2	909	100.0	1883	5	AAE27408 Human apo
3	907.4	99.8	981	10	ADC79259 Human DED
4	907.4	99.8	1230	6	ABA94362 Human APR
5	907.4	99.8	1924	6	AAD40080 Human DED
6	907.4	99.8	1924	6	AAD59062 Human DED
7	907.4	99.8	1979	12	ADQ66891 Human tum
8	907.4	99.8	2045	3	AAA95790 Apoptosis
9	907.4	99.8	2045	3	AAH3283 Human col
10	895.4	98.5	2044	3	AAH18296 Lung canc
11	825.8	90.8	1084	6	ABL38692 Human NS
12	825.8	90.8	1106	6	ABL38693 Human NS
13	799.4	87.9	1966	6	AAH59646 Human pro
14	581.2	63.9	1570	4	AAH59646 Human pro
15	541	59.5	1067	5	ADQ99183 DNA encod
16	541	59.5	1067	5	ADQ99183 DNA encod
17	541	59.5	1067	5	ADQ99183 DNA encod
18	529	58.2	1067	9	ADB48943 Novel hum
19	528	58.1	626	4	AAH07925 Human pol
20	346.2	38.1	603	5	AAF93937 Primer sp

C	21	326.6	35.9	111084	12	ADQ18808	Adq18808 Human sof
	22	301.4	33.2	303	6	AAD40075	Aad40075 Human DED
	23	301.4	33.2	303	6	AAD59057	Aad59057 Human DED
	24	193.8	21.3	1139	3	AAZ43923	Aaz43923 Human DED
	25	193.8	21.3	1830	4	AAH15765	Aah15765 Human CDN
	26	193.8	21.3	2261	4	AAH14367	Aah14367 Human CDN
	27	193.8	21.3	2261	6	ABK46133	Abk46133 CDNA enco
	28	192.2	21.1	1045	5	AAV74137	Aav74137 Human FLA
	29	192.2	21.1	1045	5	AAD43203	Aad43203 Human FLA
	30	192.2	21.1	2079	6	ABV24761	Abv24761 Human pro
	31	190.6	21.0	1142	3	AAZ43924	Aaz43924 Murine DE
	32	190.6	21.0	1200	2	AAV74138	Aav74138 Mouse FLA
	33	190.6	21.0	1200	6	AAD43204	Aad43204 Mouse FLA
	34	157.2	17.3	484	9	ACH25056	Ach25056 Human adu
	35	152.4	16.8	451	9	ACH41209	Ach41209 Human foe
C	36	119.2	13.1	769	10	ADD34431	Add34431 Mouse mit
	37	114.6	12.6	525	10	ADD34432	Add34432 Mouse mit
	38	112	12.3	553	9	ACH23692	Ach23692 Human adu
	39	111	12.2	278	11	ADT94498	Adt94498 Colon can
	40	111	12.2	422	8	ABZ18661	Abz18661 Group III
C	41	86.8	9.5	669	4	AAH05181	Aah05181 Human CDN
	42	83.8	9.2	523	4	AAH06261	Aah06261 Human CDN
	43	56.6	6.2	114955	2	AAH53491	Aax53491 Human ade
	44	54.8	6.0	980	8	ADA71128	Aad71128 Rice gene
	45	52.8	5.8	2073	8	ACC42664	Acc42664 Human LP3

ALIGNMENTS

RESULT 1	AAF27407	standard; CDNA; 909 BP.
ID	AAF27407	
XX	AAF27407;	
AC	24-APR-2001	(first entry)
XX	Human apoptosis-associated factor NT2RM1000558 partial CDNA, SEQ ID NO:1.	
DE	Human apoptosis-associated factor; NT2RM1000558; death effector domain;	
KW	DED; caspase family cleavage domain; pro-apoptotic; drug screening;	
KM	cell proliferation; ischemic disease; chronic viral disease; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	W0200104300-A1.	
XX		
PD	18-JAN-2001.	
XX		
PF	06-JUL-2000; 2000WO-JP004516.	
XX		
PR	08-JUL-1999; 99UP-00194179.	
PR	18-OCT-1999; 99US-0159586P.	
PA	(HELI-) HELIX RES INST.	
XX		
PI	Oca T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;	
XX	WPI; 2001-138348/14.	
DR	P-PSDB; AAB60386.	
XX		
PT	Polynucleotide encoding an apoptosis-associated factor protein with death	
PT	effector domain and caspase family-cleavage domain, useful in regulating	
PT	diseases with cell proliferation.	
XX		
PS	Claim 1; Page 41-43; 53pp; Japanese.	
XX		
CC	The invention relates to a novel human apoptosis-associated factor	
CC	(AAB60386, AAB60387), designated NT2RM1000558, which contains a death	
CC	effector domain (DED) and a caspase family cleavage domain and is capable	
CC	of inducing apoptosis in cells. The invention also relates to nucleic	
CC	acids encoding the protein (AAF27407, AAF27408); variants of the protein	

(particularly dominant negative variants); vectors and host cells comprising a nucleic acid which encodes an apoptosis-associated factor of the invention; the recombinant production of the protein; an antibody against the protein; and methods of screening for compounds which can regulate apoptosis. The apoptosis-related factor is useful in regulating diseases associated with cell proliferation and in screening drug candidates e.g., for regulating cell proliferation or cell death in ischemic diseases and chronic viral diseases. The present sequence represents cDNA encoding a substantial proportion of the human apoptosis-associated factor NT2RM1000558

Sequence 909 BP; 139 A; 284 C; 332 G; 154 T; 0 U; 0 Other;

Query Match 100.0%; Score 909; DB 5; Length 909;

Best Local Similarity 100.0%; Pred. No. 2.1e-171; Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGCGCTATCCGGGTCGACCCCGGCGCGTGTGGAGAGATAGTCCCTGAGCTAC 60
DB 1 ATGGCCCTATCCGGGTCGACCCCGGCGCGTGTGGAGAGATAGTCCCTGAGCTAC 60
QY 61 TACGGATGCTGTGCTTACCGTATGTTGAGAGTGGTGGCGGCAACTGACCGAGTGC 120
DB 61 TACGGATGCTGTGCTTACCGTATGTTGAGAGTGGTGGCGGCAACTGACCGAGTGC 120
QY 121 GAGCTGAGCTCTGCTCTTTCTGTGATGAGGCTCTGCGCGCCGCGAGGCTTAGCC 180
DB 121 GAGCTGAGCTCTGCTCTTTCTGTGATGAGGCTCTGCGCGCCGCGAGGCTTAGCC 180
QY 181 CGGGCCCGGAGCGGCTAGAGCTCTGCTGGAGAGTGAAGCGCGGCGAGTGGCGGAG 240
DB 181 CGGGCCCGGAGCGGCTAGAGCTCTGCTGGAGAGTGAAGCGCGGCGAGTGGCGGAG 240
QY 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGGCTGCGCCGCGCACTGCTGCCG 300
DB 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGGCTGCGCCGCGCACTGCTGCCG 300
QY 301 CACCTGCGGCGGAGCGGCGCGCGCACTGCTTCAGAACTATAGTATGACCTTCC 360
DB 301 CACCTGCGGCGGAGCGGCGCGCGCACTGCTTCAGAACTATAGTATGACCTTCC 360
QY 361 AGCTCTTCAAAAGGACAGAGGGTACGTGCGGTGCGGAGTCAAGCACTTCTGCA 420
DB 361 AGCTCTTCAAAAGGACAGAGGGTACGTGCGGTGCGGAGTCAAGCACTTCTGCA 420
QY 421 AATTCTCAGAGGAGTGTGAGAGAGAGCTCCCGCCCAACCAAGGCGCGCGGAGT 480
DB 421 AATTCTCAGAGGAGTGTGAGAGAGAGCTCCCGCCCAACCAAGGCGCGCGGAGT 480
QY 481 CGGGGCGCGGCGGAGTGTGTCAGACGCGGCGGAGAGGGGCGCCGACCGCCAG 540
DB 481 CGGGGCGCGGCGGAGTGTGTCAGACGCGGCGGAGAGGGGCGCCGACCGCCAG 540
QY 541 CAGCACTCAGAGGCGCGGCGGAGCTTCTCTGAAGGCAAAATGACCTGTGACCTCCG 600
DB 541 CAGCACTCAGAGGCGCGGCGGAGCTTCTCTGAAGGCAAAATGACCTGTGACCTCCG 600
QY 601 CGGGTTCGAGCAGAGTACTGAGCAGATGAGCAGCTTGGAGAGAGGCGGAGTCCCG 660
DB 601 CGGGTTCGAGCAGAGTACTGAGCAGATGAGCAGCTTGGAGAGAGGCGGAGTCCCG 660
QY 661 CGGCGCCGAGGCGGTGCGCGGCGGAGCTGAGCTGTGTTGGGAGGCAACCGAGTCTG 720
DB 661 CGGCGCCGAGGCGGTGCGCGGCGGAGCTGAGCTGTGTTGGGAGGCAACCGAGTCTG 720
QY 721 TCAGAGGAGCTGGGCTCTGTGTTGTGATATCAAGTTCACAGCTCTCTATCTGAGC 780
DB 721 TCAGAGGAGCTGGGCTCTGTGTTGTGATATCAAGTTCACAGCTCTCTATCTGAGC 780
QY 781 GCCTTCTGGGGGAGACTACCTGAGTGGGCGCTGCTGAGGAGCCTGCGGGGCGTGTCTG 840
DB 781 GCCTTCTGGGGGAGACTACCTGAGTGGGCGCTGCTGAGGAGCCTGCGGGGCGTGTCTG 840

```

```

QY 841 ACTGAGGCGGCTGCGGAGAGGCTGTGGGCGGAGAGGCTGTCCGCTGCTGATGATGAGT 900
DB 841 ACTGAGGCGGCTGCGGAGAGGCTGTGGGCGGAGAGGCTGTCCGCTGCTGATGATGAGT 900
QY 901 GAGGCTGAC 909
DB 901 GAGGCTGAC 909

```

RESULT 2

AAE27408
ID AAE27408 standard; cDNA; 1883 BP.

AAE27408;

24-APR-2001 (first entry)

Human apoptosis-associated factor NT2RM1000558 cDNA, SEQ ID NO:3.

Human; apoptosis-associated factor; NT2RM1000558; death effector domain;

DED; caspase family cleavage domain; pro-apoptotic; drug screening;

cell proliferation; ischemic disease; chronic viral disease; ss.

Homo sapiens.

WO200104300-A1.

06-JUL-2000; 2000MO-JP004516.

08-JUL-1999; 99JP-00194179.

18-OCT-1999; 99US-0159586P.

(HELI-) HELIX RES INST.

Ota T, Isegai T, Nishikawa T, Kawai Y, Miyoshi S, Sato S;

WPI; 2001-138348/14.

P-PSDB; AAB60387.

Polynucleotide encoding an apoptosis-associated factor protein with death

effector domain and caspase family-cleavage domain, useful in regulating

diseases with cell proliferation.

Claim 3; Page 44-47; 53p; Japanese.

The invention relates to a novel human apoptosis-associated factor

(AAB60386, AAB60387), designated NT2RM1000558, which contains a death

effector domain (DED) and a caspase family cleavage domain and is capable

of inducing apoptosis in cells. The invention also relates to nucleic

acids encoding the protein (AAE27407, AAE27408); variants of the protein

(particularly dominant negative variants); vectors and host cells

comprising a nucleic acid which encodes an apoptosis-associated factor

of the invention; the recombinant production of the protein; an antibody

against the protein; and methods of screening for compounds which can

regulate apoptosis. The apoptosis-related factor is useful in regulating

diseases associated with cell proliferation and in screening drug

candidates e.g., for regulating cell proliferation or cell death in

ischemic diseases and chronic viral diseases. The present sequence

represents a full-length cDNA encoding the human apoptosis-associated

factor NT2RM1000558

Sequence 1883 BP; 343 A; 595 C; 588 G; 357 T; 0 U; 0 Other;

Query Match 100.0%; Score 909; DB 5; Length 1883;

Best Local Similarity 100.0%; Pred. No. 2.2e-171; Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGCCCTATCCGGGTCGACCCCGGCGCGTGTGGAGAGATAGTCCCTGAGCTAC 60
DB 124 ATGGCCCTATCCGGGTCGACCCCGGCGCGTGTGGAGAGATAGTCCCTGAGCTAC 183

```



```

QY 61 TACGGATGCTGTGCTTACCGTATGTTCAAGGTGTGGCGGGCAATGACCGAGTGC 120
DB 184 TACGGAGTGTGCTGTACCGTATGTTCAAGGTGTGGCGGGCAATGACCGAGTGC 243
QY 121 GAGCTGAGGCTCTGGGCTTTCTGTGTGATGAGCTCTGTGGCGCGCGGAGGCTTAGCC 180
DB 244 GAGCTGAGGCTCTGGGCTTTCTGTGTGATGAGCTCTGTGGCGCGCGGAGGCTTAGCC 303
QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGTGTGAGTGTGAGCGCGCGGAGGCTTAGCC 240
DB 304 CGGGCCCGCAGCGGCTAGAGCTCTGTGTGAGTGTGAGCGCGCGGAGGCTTAGCC 363
QY 241 AGCAACTGCGGCTGTGGGCAACTCTGTGGCGGTGTGGCGCGCGGAGGCTTAGCC 300
DB 364 AGCAACTGCGGCTGTGGGCAACTCTGTGGCGGTGTGGCGCGCGGAGGCTTAGCC 423
QY 301 CACCTGGCGCGCAGCGGCGCGGCAAGTGTCTCCAGAACGCTATAGTATGAGCACTCC 360
DB 424 CACCTGGCGCGCAGCGGCGCGGCAAGTGTCTCCAGAACGCTATAGTATGAGCACTCC 483
QY 361 AGCTCTTCAAGAGAGACAGAGGTAGTGTGCGCGCTGTGGCAAGTGTGCA 420
DB 484 AGCTCTTCAAGAGAGACAGAGGTAGTGTGCGCGCTGTGGCAAGTGTGCA 543
QY 421 AATTCTCAGCAGGGTCAGTGGGAGACAGGCTTCCCCCAACCAAGCGGAGCGAGT 480
DB 544 AATTCTCAGCAGGGTCAGTGGGAGACAGGCTTCCCCCAACCAAGCGGAGCGAGT 603
QY 481 CGGGGCGCGCGCAGTGTGTGTGCGAGAGGGGCGGAGAGGGGCGCGCGGAGCGGAG 540
DB 604 CGGGGCGCGCGCAGTGTGTGTGCGAGAGGGGCGGAGAGGGGCGCGCGGAGCGGAG 663
QY 541 CAGCAGTACAGAGCGCGCAGACCTTCTGTGAGGCAAGTGTGAGTGTGAGTGTG 600
DB 664 CAGCAGTACAGAGCGCGCAGACCTTCTGTGAGGCAAGTGTGAGTGTGAGTGTG 723
QY 601 CGGGTTCAGAGAGTACTGTGCGAGTGTGCGAGCTTGTGAGAGAGGCGGTGAGT 660
DB 724 CGGGTTCAGAGAGTACTGTGCGAGTGTGCGAGCTTGTGAGAGAGGCGGTGAGT 783
QY 661 CGGGCCCGCAGGCTGTGGCGCGGAGTGTGTTGTGGGCAAGCGGAGTGTGCGC 720
DB 784 CGGGCCCGCAGGCTGTGGCGCGGAGTGTGTTGTGGGCAAGCGGAGTGTGCGC 843
QY 721 TCAAGGAGCTGTGGCTGTGTTGTGATCAATCAAGTCTCAGAGCTCTCCATCTGAGC 780
DB 844 TCAAGGAGCTGTGGCTGTGTTGTGATCAATCAAGTCTCAGAGCTCTCCATCTGAGC 903
QY 781 GCTTCTGTGGGCGACTACTAGTGTGCGCGCTGTGCAAGCGCTGTGGGAGGCTTCTG 840
DB 904 GCTTCTGTGGGCGACTACTAGTGTGCGCGCTGTGCAAGCGCTGTGGGAGGCTTCTG 963
QY 841 ACTAGAGCGCTGTGAGAGAGGCTGTGGCGGAGGCTGTGGCGGCTGTGAGTGTGAT 900
DB 964 ACTAGAGCGCTGTGAGAGAGGCTGTGGCGGAGGCTGTGGCGGCTGTGAGTGTGAT 1023
QY 901 GAGGCTGAC 909
DB 1024 GAGGCTGAC 1032

```

RESULT 3
 ADCT9259
 ID ADCT9259 standard; cDNA; 981 BP.

AC ADCT9259;

DT 01-JAN-2004 (first entry)

XX Human DEDD2 encoding cDNA SEQ ID NO:1.

XX human; death effector domains containing DNA-binding protein;

KW DED-containing DNA-binding protein; DEDD2; cell death; gene therapy;

```

KW cytosolic; cancer; chronic myeloid leukaemia; gene; ss.
XX Homo sapiens.
OS
XX Key location/Qualifiers
FT CDS 1..981
FT /tag= a
FT /product= "DEDD2"
XX MO2003054195-A1.
XX 03-JUL-2003.
XX 20-DEC-2002; 2002WO-JP013371.
XX PF
XX 20-DEC-2001; 2001JP-00387854.
XX PR
XX 18-JUL-2002; 2002JP-00209458.
XX PR
XX (MORG ) MORINAGA MILK IND CO LTD.
XX PA
XX Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX WPI; 2003-569246/53.
XX DR
XX P-PSDB; ADCT9260.
XX PT
XX DNA encoding cell death proteins for treatment of kidney, large intestine
XX and prostate cancers and leukemia.
XX PS
XX Claim 2; Page 18-20; 26pp; Japanese.
XX CC
XX The present sequence encodes a human death effector domains (DED)
XX containing DNA-binding protein (DEDD) protein, designated DEDD2, that
XX causes cell death. Also described: (1) primer and probe for investigation
XX of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
XX cytosolic activity. DEDD2 can be used in the diagnosis and treatment of
XX cancers of the kidney, large intestine and prostate, and acute and
XX chronic myeloid leukaemia.
XX SQ
XX Sequence 981 BP; 151 A; 303 C; 363 G; 164 T; 0 U; 0 Other;
XX
XX Query Match 99.8%; Score 907.4; DB 10; Length 981;
XX Best Local Similarity 99.9%; Pred. No. 4.4e-171;
XX Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGTATCCGGGTGAGACCCCGCGGCTGTGGAGAGAGTATGCTGAGACTAC 60
DB 1 ATGGGCGTATCCGGGTGAGACCCCGCGGCTGTGGAGAGAGTATGCTGAGACTAC 60
QY 61 TACGGAGTGTGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
DB 61 TACGGAGTGTGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
QY 61 TACGGAGTGTGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
DB 61 TACGGAGTGTGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
QY 121 GAGCTGAGGCTCTGGGCTTTCTGTGTGATGAGGCTCTGTGGCGCGCGGAGGCTTAGCC 180
DB 121 GAGCTGAGGCTCTGGGCTTTCTGTGTGATGAGGCTCTGTGGCGCGCGGAGGCTTAGCC 180
QY 121 GAGCTGAGGCTCTGGGCTTTCTGTGTGATGAGGCTCTGTGGCGCGCGGAGGCTTAGCC 180
DB 121 GAGCTGAGGCTCTGGGCTTTCTGTGTGATGAGGCTCTGTGGCGCGCGGAGGCTTAGCC 180
QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGTGTGAGTGTGAGCGCGGAGGCTTAGCC 240
DB 181 CGGGCCCGCAGCGGCTAGAGCTCTGTGTGAGTGTGAGCGCGGAGGCTTAGCC 240
QY 241 AGCAACTGCGGCTGTGGGCAACTCTGTGGCGGTGTGGCGCGCGGAGGCTTAGCC 300
DB 241 AGCAACTGCGGCTGTGGGCAACTCTGTGGCGGTGTGGCGCGCGGAGGCTTAGCC 300
QY 301 CACCTGGCGCGCAGCGGCGCGGCAAGTGTCTCCAGAACGCTATAGTATGAGCACTCC 360
DB 301 CACCTGGCGCGCAGCGGCGCGGCAAGTGTCTCCAGAACGCTATAGTATGAGCACTCC 360
QY 361 AGCTCTTCAAGAGAGACAGAGGTAGTGTGCGCGCTGTGGCAAGTGTGCA 420
DB 361 AGCTCTTCAAGAGAGACAGAGGTAGTGTGCGCGCTGTGGCAAGTGTGCA 420
QY 421 AATTCTCAGCAGGGTCAGTGGGAGACAGGCTTCCCCCAACCAAGCGGAGCGGAGT 480

```

```

Db      421 AATTCTCAGAGGAGTCTGAGAGACAGGCTCCCCCAACAGAGCGGCGAGT 480
QY      481 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGAGGCGCCAGACCCAG 540
Db      481 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGAGGCGCCAGACCCAG 540
QY      541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAAGTACCTGTACATCCGCTC 600
Db      541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAAGTACCTGTACATCCGCTC 600
QY      601 CGGGTTCAGACAGAGTACTGCGAGCATGGGCGAGCCTTGGAGAGGCGTGGATCCCG 660
Db      601 CGGGTTCAGACAGAGTACTGCGAGCATGGGCGAGCCTTGGAGAGGCGTGGATCCCG 660
QY      661 CGGCCCCGAGCGCTGCGCGGCACTGAGACGTGTTTGGGAGGCAACGAGTGTGCG 720
Db      661 CGGCCCCGAGCGCTGCGCGGCACTGAGACGTGTTTGGGAGGCAACGAGTGTGCG 720
QY      721 TCAAGGAGACTGGGCTCTGTGTGTTTGTGACATCAAGTTCTCAGACTCTCTATCTG 780
Db      721 TCAAGGAGACTGGGCTCTGTGTGTTTGTGACATCAAGTTCTCAGACTCTCTATCTG 780
QY      781 GCCTTCGGGGGCACTACCTGAGTGGGCGCCTGTGAGGCGCTGGGGGGGCTTCTG 840
Db      781 GCCTTCGGGGGCACTACCTGAGTGGGCGCCTGTGAGGCGCTGGGGGGGCTTCTG 840
QY      841 ACTGAGCCCTGGAGAGAGCTGTGTGGGCGGAGAGGCTTGTGCTGTGTCAGTGTGAT 900
Db      841 ACTGAGCCCTGGAGAGAGCTGTGTGGGCGGAGAGGCTTGTGCTGTGTCAGTGTGAT 900
QY      901 GAGGCTGAC 909
Db      901 GAGGCTGAC 909

RESULT 4
ABA94362
ID      ABA94362 standard; DNA; 1230 BP.
XX
AC      ABA94362;
XX
DT      26-MAR-2002 (first entry)
XX
DE      Human APRG polypeptide (Incyte ID. 3102521CD1) encoding cDNA.
XX
KM      APRG; apoptosis regulator; cytosstatic; antiatherosclerotic; osteopathic;
KM      antiarteriosclerotic; hepatotropic; antiproliferative; anthelmintic; human;
KM      antiatherogenic; antineuritic; antiasthmatic; antithyroid; anti-HIV; cancer;
KM      antiinflammatory; antidiabetic; antitumor; nephrotropic; ophthalmological;
KM      immunosuppressive; dermatological; antitumor; antirheumatic; fungicide;
KM      antitubercular; antibacterial; virucide; antiparasitic; protozoacide;
KM      tranquilizer; vulnerary; gynecological; vasotropic; gene therapy; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      CDS
FT      127..1107
FT      /*tag= a
FT      /product= "APRG polypeptide"
XX
XX      MO200192527-A2.
XX
XX      06-DEC-2001.
XX
XX      30-MAY-2001; 2001MO-US017581.
XX
XX      01-JUN-2000; 2000US-0209407P.
XX      30-NOV-2000; 2000US-0250326P.
XX
XX      (INCY-) INCYTE GENOMICS INC.
XX
XX      Tang YT, Azimzai Y, Yue H, Burford N, Ding L, Elliott VS;
XX
XX      PI
```

```

PI      Patterson C, Baughn MR;
XX
XX      WPI; 2002-114350/15.
DR      P-PSDB; ABB07263.
XX
XX      Novel human apoptosis regulator polypeptides and polynucleotides for
PT      diagnosing, preventing, treating cell proliferative, immunological and
PT      reproductive disorders and for identifying modulators of therapeutic use.
XX
XX      Claim 5; Page 101; 103pp; English.
XX
XX      The invention provides human apoptosis regulator (APRG) polypeptides and
CC      polynucleotides. The APRG polypeptides, polynucleotides and modulators
CC      are useful for diagnosis, treatment and prevention of cell proliferative,
CC      immunological and reproductive disorders. The cell proliferative
CC      disorders include cancer, actinic keratosis, arteriosclerosis,
CC      atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, and
CC      immunological disorders include acquired immunodeficiency syndrome
CC      (AIDS), adult respiratory distress syndrome, Addison's disease,
CC      ankylosing spondylitis, amyloidosis, allergies, anemia, osteoporosis,
CC      autoimmune hemolytic anemia, asthma, autoimmune thyroiditis, Crohn's
CC      disease, contact dermatitis, diabetes mellitus, gout, Graves' disease,
CC      glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
CC      erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis,
CC      ureitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC      infections and trauma. Reproductive disorders include disorders of
CC      prolectin production, infertility, endometriosis, polycystic ovary
CC      syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,
CC      disruptions of spermatogenesis, cancer of testis and prostate, impotence,
CC      carcinoma of male breast and gynecomastia. The APRG polynucleotides are
CC      useful for creating knockin humanized animals or transgenic animals to
CC      model human disease and to detect and quantify gene expression in
CC      biopsied tissues in which expression of APRG is correlated with disease.
CC      APRG, fragments of it and antibodies specific for APRG are useful as
CC      elements on a microarray which is useful to monitor or measure protein-
CC      protein interactions, drug-target interactions and gene expression
CC      profiles. The present sequence represents a human APRG polypeptide
CC      encoding cDNA
XX
XX      Sequence 1230 BP; 205 A; 382 C; 433 G; 210 T; 0 U; 0 Other;
SQ
Query Match          99.8%; Score 907.4; DB 6; Length 1230;
Best Local Similarity 99.9%; Pred. No. 4.4e-171;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 ATGGCGCTATCCCGGATGCAACCCCGGCGCTGTGGAGAGAGATGCTGTGACTAC 60
Db      127 ATGGCGCTATCCCGGATGCAACCCCGGCGCTGTGGAGAGAGATGCTGTGACTAC 186
QY      61 TACGGATGCTGTGCTTCAACCTATGTTGAGAGTGTGGGCGGCACTGACCGAGTGC 120
Db      187 TACGGATGCTGTGCTTCAACCTATGTTGAGAGTGTGGGCGGCACTGACCGAGTGC 246
QY      121 GAGCTGAGCTCTGCGCTTTCGTGATGAGGCTCTGCGGCGCGCGGAGCTTAGCC 180
Db      247 GAGCTGAGCTCTGCGCTTTCGTGATGAGGCTCTGCGGCGCGCGGAGCTTAGCC 306
QY      181 CGGGCCCGAGCGGCTTAGAGCTCTGTGAGCTGAGAGCGCGCGGAGAGTGTGCGAG 240
Db      307 CGGGCCCGAGCGGCTTAGAGCTCTGTGAGCTGAGAGCGCGCGGAGAGTGTGCGAG 366
QY      241 AGCAACTGGGCTGTGTTGGGCAACTCTGCGGCTGTGAGCGCGCGCACTGTGCGG 300
Db      367 AGCAACTGGGCTGTGTTGGGCAACTCTGCGGCTGTGAGCGCGCGCACTGTGCGG 426
QY      301 CACCTGGCGGCAAGAGGCGCGCGCGCAAGTCTCCGAAAGCTATGACTATGCACTCC 360
Db      427 CACCTGGCGGCAAGAGGCGCGCGCGCAAGTCTCCGAAAGCTATGACTATGCACTCC 486
QY      361 AGCTTTCAAGAGAGAGAGAGGTAGCTGCGGCTGTGCGAGTCAAGCACTTGTGCA 420
Db      487 AGCTTTCAAGAGAGAGAGGTAGCTGCGGCTGTGCGAGTCAAGCACTTGTGCA 546
```

```

QY 421 AATTCTCAGCAGGGTCACTGGGAGACAGGCTTCCCCCAACCAACGGCAGCGCGAGT 480
D 547 AATTCTCAGCAGGGTCACTGGGAGACAGGCTTCCCCCAACCAACGGCAGCGCGAGT 606
QY 481 CCGGGCCCGCCAGTGGTGGTCCAGAGCGGGCGGAGAGGGGCGCCAGCCCGAG 540
D 607 CCGGGCCCGCCAGTGGTGGTCCAGAGCGGGCGGAGAGGGGCGCCAGCCCGAG 666
QY 541 CAGAGTCAGAGCCCGCAGACCTTCTCTGAGAGCAAGTGAACCTGTGATCCGGCTC 600
D 667 CAGAGTCAGAGCCCGCAGACCTTCTCTGAGAGCAAGTGAACCTGTGATCCGGCTC 726
QY 601 CCGGTTGAGCAGAGTACTGCGAGCAGTGGCCAGCTTTGAGAGCAGGCGTGGCATCCGG 660
D 727 CCGGTTGAGCAGAGTACTGCGAGCAGTGGCCAGCTTTGAGAGCAGGCGTGGCATCCGG 786
QY 661 CCGGCCAGAGCGCTGGCGCGCAGCTGACGTTGTTGGGCGAGCCACCGCAGTGTGGC 720
D 787 CCGGCCAGAGCGCTGGCGCGCAGCTGACGTTGTTGGGCGAGCCACCGCAGTGTGGC 846
QY 721 TCAAGGACCTGGGCTGTGTGTTGTGACATCAAGTCTCAGAGCTCTCCATCTGGAC 780
D 847 TCAAGGACCTGGGCTGTGTGTTGTGACATCAAGTCTCAGAGCTCTCCATCTGGAC 906
QY 781 GCTTTCTGGGGCGACTACCTGAGTGGCGCCCTGCTGACAGGCTTGGCGGGCGTGTCTG 840
D 907 GCTTTCTGGGGCGACTACCTGAGTGGCGCCCTGCTGACAGGCTTGGCGGGCGTGTCTG 966
QY 841 ACTAGAGCCCTGCGAGAGGCTGTGGGCGCGAGGCTGTTCCTCTGCTGATAGTGTGAT 900
D 967 ACTAGAGCCCTGCGAGAGGCTGTGGGCGCGAGGCTGTTCCTCTGCTGATAGTGTGAT 1026
QY 901 GAGGCTGAC 909
D 1027 GAGGCTGAC 1035

RESULT 5
AAD40080
ID AAD40080 standard; DNA; 1924 BP.
XX
AC AAD40080;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human DED4 (death effector domain) gene.
XX
KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
KW immunosuppressive; gene therapy; antisense therapy; gene; de.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 91..1044
FT FT /*tag= a
FT FT /product= "Human DED4"
FT FT /note= "No stop codon"
FT FT /partial
FT misc_feature 157..222
FT FT /*tag= b
FT FT /note= "Nuclear localisation sequence"
XX
FN WO200240680-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001WO-US044844.
XX
PR 17-NOV-2000; 2000US-00715893.
PR 29-JUN-2001; 2001US-0301889P.

```

```

XX (BURN-) BURNHAM INST.
PA
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
PI Steiner-Liwen F;
XX WPI; 2002-500222/53.
DR P-PSDB; AAE24860.
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
XX Claim 19; Page 184-186; 209pp; English.
XX
CC The invention relates to an isolated polypeptide comprising a death
CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
CC is useful for identifying a binding agent, preferably a protein or a drug
CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
CC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
CC NIPD (NGFR-interacting Death Domain), with a candidate binding agent and
CC detecting the association of the domain and the candidate binding agent,
CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
CC spectroscopy (MS) and FPA. The invention is useful for modulating the
CC level of a cell process such as cell proliferation, cell adhesion, cell
CC stress responses, responses to microbial infection and B cell
CC immunoglobulin class switching, in particular apoptosis within a cell.
CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
CC CTDD DD protein is useful for detecting a Chlamydia infection. The
CC invention is useful for modulating the activity of oncogenic proteins,
CC for treating a pathology caused by the oncogenic proteins and for
CC treating bacterial infections by modulating the activity of bacterial
CC proteins. The protein and antibody specific for it are useful for
CC discovery of drugs that suppress infection, inflammation, allergy,
CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
CC is useful for treating immune-based pathologies, pathologies associated
CC with cell division, inflammatory diseases such as sepsis, fibrosis,
CC arthritis, graft versus host disease. The invention is used in antisense
CC therapy and gene therapy. The present sequence is human DED4 gene
XX
XX
SQ Sequence 1924 BP; 360 A; 599 C; 594 G; 368 T; 0 U; 3 Other;
XX
Query Match 99.8%; Score 907.4; DB 6; Length 1924;
Best Local Similarity 99.9%; Pred. No. 4.6e-111;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTGCCTGACTAC 60
D 91 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTGCCTGACTAC 150
QY 61 TACGGATGCTGTGCTTACCGGTATGTTGAGGTGTGGCGGGCAACTGACGAGTGC 120
D 151 TACGGATGCTGTGCTTACCGGTATGTTGAGGTGTGGCGGGCAACTGACGAGTGC 210
QY 121 GAGCTGAGAGCTCCGGGCTTCTGCTGGATGAGGCTCTCGGCGCGCGGAGCTTACC 180
D 211 GAGCTGAGAGCTCCGGGCTTCTGCTGGATGAGGCTCTCGGCGCGCGGAGCTTACC 270
QY 181 CCGGCCCGCAGCGGCTTGAAGCTTCTGAGAGCTGAGCGCCCGGAGCACTGCGCGAG 240
D 271 CCGGCCCGCAGCGGCTTGAAGCTTCTGAGAGCTGAGCGCCCGGAGCACTGCGCGAG 330
QY 241 AGCAACTGGGCTGTGGGGCACTCTGTGCGCTGTGCGCGCCGACAGACTCTGCGG 300
D 331 AGCAACTGGGCTGTGGGGCACTCTGTGCGCTGTGCGCGCCGACAGACTCTGCGG 390
QY 301 CACCTGGGCGCAGAGCGCGCGCGCAGTGTCTGCAGAAAGCTATTAGCACTCC 360
D 391 CACCTGGGCGCAGAGCGCGCGCGCAGTGTCTGCAGAAAGCTATTAGCACTCC 450
QY 361 AGCTCTTCAAGAAGACAGAGGATAGCTGCCGTGCCGTGGCAGTCAAGCAGTTCTGCA 420

```

```

Db      451 AGCTCTTCAAGAGGAGCAGAGGGTAGCTGCGCTGCGGTGAGGAGTCAAGCACTTCTGCA
Qy      421 AATTCTCAGCAGGAGTCACTGAGAGCAGAGCTCCCTCCCAACCAAGCGGCGAGGAGT 480
Db      511 AATTCTCAGCAGGAGTCACTGAGAGCAGAGCTCCCTCCCAACCAAGCGGCGAGGAGT 570
Qy      481 CGGGGCGGCGCAGTGTGTGTGTCAGACGGGCGGAGAGAGGGGCCCGGACCGGACCCGAG 540
Db      571 CGGGGCGGCGCAGTGTGTGTGTCAGACGGGCGGAGAGAGGGGCCCGGACCGGACCCGAG 630
Qy      541 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAGGAAAGTGAACCTGTGACATCCGGCTC 600
Db      631 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAGGAAAGTGAACCTGTGACATCCGGCTC 690
Qy      601 CGGGTTCAGACAGTACTCCGAGCATGAGGACCGCTTGAAGAGAGGGCGTGGCATCCCG 660
Db      691 CGGGTTCAGACAGTACTCCGAGCATGAGGACCGCTTGAAGAGAGGGCGTGGCATCCCG 750
Qy      661 CGGCCCCAGGCGCTGCGCGGCGAGCTGAGACGTGTGTTGGGAGGCGCACCGAGTGTGCGC 720
Db      751 CGGCCCCAGGCGCTGCGCGGCGAGCTGAGACGTGTGTTGGGAGGCGCACCGAGTGTGCGC 810
Qy      721 TCAAGGAGACCTGAGGCTGTGTGTTGTGAGATCAAGTTCTCAGAGCTTCTCTATCTGGAC 780
Db      811 TCAAGGAGACCTGAGGCTGTGTGTTGTGAGATCAAGTTCTCAGAGCTTCTCTATCTGGAC 870
Qy      781 GCCTTCGAGGAGCACTACTGAGTGGGCGCCTGCTGAGGACCGCTGAGGAGCGAGTTCCTG 840
Db      871 GCCTTCGAGGAGCACTACTGAGTGGGCGCCTGCTGAGGACCGCTGAGGAGCGAGTTCCTG 930
Qy      841 ACTGAGCGCTGCGAGAGGCTGTGTGGGCGGAGAGGCTGTGCTGCTGTGATGTGTGAT 900
Db      931 ACTGAGCGCTGCGAGAGGCTGTGTGGGCGGAGAGGCTGTGCTGCTGTGATGTGTGAT 990
Qy      901 GAGGCTGAC 909
Db      991 GAGGCTGAC 999

```

RESULT 6
AADS9062 standard; cDNA; 1924 BP.

AC AADS9062;
DT 18-DEC-2003 (first entry)
DE Human DED4 full length gene.

XX Human; death Domain; DD; death effector domain; DED; cell proliferation;
XX Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;
KM neural growth factor receptor-interacting death domain; cell adhesion;
KM vasotropic; microbial infection; inflammation; allograft rejection; CTDP;
KM cell stress response; benign prostatic hyperplasia; antibacterial; NIDP;
KM apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;
KM neoplasia; restenosis; immunosuppressive; antibody therapy; cytoskeletal;
KM keloid; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 91..1044
FT /*tag= a
FT /*product= "Human DED4 protein"
FT /*note= "No stop codon"

FT misc_signal
FT 157..222
FT /*tag= b
FT /*note= "Nuclear localisation sequence"

XX US2003049702-A1.
XX

```

PD      13-MAR-2003.
XX
XX      15-NOV-2001; 2001US-00001254.
PF
XX      17-NOV-2000; 2000US-00715893.
XX
PR      17-NOV-2000; 2000US-0367360P.
PR      29-JUN-2001; 2001US-0301889P.
XX
XX      (REED/) REED J C.
PA      (GODZ/) GODZIK A.
PA      (PAWL/) PAWLOWSKI K.
PA      (PIOR/) FIORENTINO L.
PA      (LEES/) LEE S H.
PA      (ROTH/) ROTH W.
PA      (STEN/) STENNER-LIEWEN F.
XX
PI      Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
PI      Stenner-Liewen F;
DR      WPI; 2002-500222/53.
DR      P-PSDB; AAE38903.
XX
PT      New polypeptide comprising a death domain or death effector domain,
PT      useful for discovery of drugs that suppress infection, inflammation,
PT      allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
XX      Claim 19; Page 42-44; 99pp; English.
XX
XX      The present invention provides novel death domain (DD) and death effector
XX      domain (DED) proteins and nucleic acids encoding them. The invention also
XX      provides death domain containing protein such as Chlamydia trachomatis
XX      death domain containing protein (CTDD) DD and neural growth factor
XX      receptor-interacting death domain (NIDP) DD. The invention is useful for
XX      identifying a binding agent (e.g. protein or drug) that binds a DD, DED
XX      or NB-ARC domain from DAP3, IRAK4, CTDP, DED4 or NIDP with a candidate
XX      binding agent and identifying an effective agent (e.g. protein or drug)
XX      that modulates the association of a DD, DED or NB-ARC domain with protein
XX      that binds the DD, DED or NB-ARC domain. The invention is also useful for
XX      modulating the level of cell process such as apoptosis, cell adhesion,
XX      cell proliferation, cell stress responses, responses to microbial
XX      infection and B cell immunoglobulin class switching. DPs, DEDs and NB-ARC
XX      domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are
XX      useful for discovery of drugs that suppress infection, autoimmunity,
XX      inflammation, allergy, allograft rejection, sepsis and other diseases.
XX      DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
XX      autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
XX      hyperplasia, neoplasia, keloid, benign prostatic hyperplasia, fibrosis,
XX      inflammatory hyperplasia and smooth muscle cell proliferation in arteries
XX      following balloon angioplasty (restenosis). The invention is also used in
XX      antibody therapy and gene therapy. The present sequence is human DED4
XX      full length gene. The DED4 gene is located on chromosome 19
XX
SQ      Sequence 1924 BP; 360 A; 599 C; 594 G; 368 T; 0 U; 3 Other;

```

Query Match 99.8%; Score 907.4; DB 6; Length 1924;
Best Local Similarity 99.9%; Pred. No. 4,6e-171;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 ATGGCGCTATCGGGTGCAGACCCCGGCGCGTGTGGAGAGAGATGCTGTGACTAC 60
Db      91 ATGGCGCTATCGGGTGCAGACCCCGGCGCGTGTGGAGAGAGATGCTGTGACTAC 150
Qy      61 TACGGATGCTGTGCTTCACTTATGATGAGTGTGGGCGGCAACTGACGAGTGC 120
Db      151 TACGGATGCTGTGCTTCACTTATGATGAGTGTGGGCGGCAACTGACGAGTGC 210
Qy      121 GAGCTGAGGCTCTGCGCTTTTCTGCTGATGAGGCTTCGCGCGCGGAGGCTTAGCC 180
Db      211 GAGCTGAGGCTCTGCGCTTTTCTGCTGATGAGGCTTCGCGCGCGGAGGCTTAGCC 270
Qy      181 CGGGCCGAGCGGCGCTAGAGCTCTGAGAGCTGAGAGCGCGGCGGAGCTGGGAG 240
Db      271 CGGGCCGAGCGGCGCTAGAGCTCTGAGAGCTGAGAGCGCGGCGGAGCTGGGAG 330

```


QY 241 AGCAACCTGGGCTGCTGGGGCAACTCTCTGCGGTGCTGGCCCGCCACGACTGTGCTCCG 300
 DB 401 AGCAACCTGGGCTGCTGGGGCAACTCTCTGCGGTGCTGGCCCGCCACGACTGTGCTCCG 460
 QY 301 CACCTGGCGGCAAGCGGGCGCGGCGGAGTCTTCAGAAAGCTATGACTATGACACTCC 360
 DB 461 CACCTGGCGGCAAGCGGGCGCGGCGGAGTCTTCAGAAAGCTATGACTATGACACTCC 520
 QY 361 AGCTCTTCAAAAGAGACAGAGGGTAGTGTGCTGCGCGCTCGGAGTCAAGCACTTGTCA 420
 DB 521 AGCTCTTCAAAAGAGACAGAGGGTAGTGTGCTGCGCGCTCGGAGTCAAGCACTTGTCA 580
 QY 421 AATTCTCAGAGGGTCACTGGGAGACAGGCTTCCCTCCCAACCAAGCGGCGCGGAGT 480
 DB 581 AATTCTCAGAGGGTCACTGGGAGACAGGCTTCCCTCCCAACCAAGCGGCGCGGAGT 640
 QY 481 CGGGGGCGGGCCAGTGTGTGTCAGACGGGGCGGAGAGGGGGCCCGAGCGGACCCGAG 540
 DB 641 CGGGGGCGGGCCAGTGTGTGTCAGACGGGGCGGAGAGGGGGCCCGAGCGGACCCGAG 700
 QY 541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAAAGTGACTGTGACATCCGCTC 600
 DB 701 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAAAGTGACTGTGACATCCGCTC 760
 QY 601 CGGGTTTGAGAGAGTATCTGCGAGCATGGGCGAGCTTGTGAGAGAGGGGCTGGCATCCGG 660
 DB 761 CGGGTTTGAGAGAGTATCTGCGAGCATGGGCGAGCTTGTGAGAGAGGGGCTGGCATCCGG 820
 QY 661 CGGGCCCGAGGCGCTGGGCGGCGGCGGCTGGAAGTGTGTTGGGAGGCAACGAGTGTGCGG 720
 DB 821 CGGGCCCGAGGCGCTGGGCGGCGGCGGCTGGAAGTGTGTTGGGAGGCAACGAGTGTGCGG 880
 QY 721 TCAAGGACCTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGAC 780
 DB 881 TCAAGGACCTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGAC 940
 QY 781 GCCTTCTGGGGCGACTATCTGAGTGGGCGCTGTGACAGGCGCTGGGGGCGTGTCTTG 840
 DB 941 GCCTTCTGGGGCGACTATCTGAGTGGGCGCTGTGACAGGCGCTGGGGGCGTGTCTTG 1000
 QY 841 ACTGAGGCGCTGGGAGAGGCTGTGGGCGGGAAGGCTGTGGCTGCTGATCAAGTGGAT 900
 DB 1001 ACTGAGGCGCTGGGAGAGGCTGTGGGCGGGAAGGCTGTGGCTGCTGATCAAGTGGAT 1060
 QY 901 GAGGCTGAC 909
 DB 1061 GAGGCTGAC 1069

RESULT 8
 ID AAA95790 standard; cDNA, 2045 BP.
 AC AAA95790;

28-FEB-2001 (first entry)

Apoptosis related gene 1 clone HUDOK36.

KM Cytostatic; antiinflammatory; immunosuppressive; antisclerotic; cardiant;
 KM viral; anti-AIDS; vasodilator; anti-ischaemic; antiparkinsonian; ss;
 KM anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;
 KM colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;
 KM autoimmune disorder; multiple sclerosis; viral infection.

OS Homo sapiens.

PN MO200056752-A2.

PD 28-SEP-2000.

PF 15-MAR-2000; 2000MO-US006642.

XX 24-MAR-1999; 99US-0126018P.
 PR 17-JUN-1999; 99US-0139638P.
 PR 18-AUG-1999; 99US-0149449P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ruben SM, Ni J, Young PA;
 XX
 XX WPI; 2000-587660/55.
 DR P-PSDB; AAB15551.
 XX
 PT Nucleic acids encoding human apoptosis associated protein, useful for the
 PT prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's
 PT disease, inflammation and ischemic injury.
 PS
 PS Claim 1; Page 244-245; 273pp; English.
 XX
 CC The invention relates to the isolation of genes encoding 9 human
 CC apoptosis-related proteins. The nucleotide sequences AAA95790-A95798
 CC encode the human apoptosis related proteins AAB15551-B15559. The genes
 CC can be used to generate fusion proteins by linking to the gene for the
 CC human immunoglobulin G Fc (196 Fc) portion (AAA95799) for increasing the
 CC stability of the fusion protein as compared to the human protein only.
 CC The gene and encoded protein may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate apoptosis associated
 CC protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,
 CC melanomas and lymphomas), inflammation, autoimmune disorders (e.g.
 CC multiple sclerosis) and viral infections (e.g. herpes)
 XX
 SQ Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;
 Query Match 99.8%; Score 907.4; DB 3; Length 2045;
 Best Local Similarity 99.9%; Pred. No. 4, 6e-171;
 Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGGCCCTATCCGGGTGACCCCGGCGCTGTGGAGAGATAGTACCTGACTAC 60
 DB 121 ATGGCCCTATCCGGGTGACCCCGGCGCTGTGGAGAGATAGTACCTGACTAC 180
 QY 61 TACGGATGCTGTGCTTCACTGATGTTGAGTGTGGCGGGAACCTGACGAGTGC 120
 DB 181 TACGGATGCTGTGCTTCACTGATGTTGAGTGTGGCGGGAACCTGACGAGTGC 240
 QY 121 GAGCTGAGCTCTGACCTTTTGTGTGATGAGCTTCTGCGCGCCGAGGCTTAGCC 180
 DB 241 GAGCTGAGCTCTGACCTTTTGTGTGATGAGCTTCTGCGCGCCGAGGCTTAGCC 300
 QY 181 CGGGCCCGAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGGCGAGTGGCGAG 240
 DB 301 CGGGCCCGAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGGCGAGTGGCGAG 360
 QY 241 AGCAACCTGGGCTGTGCTGGGCAACTCTGCGGCTGTGGCCCGCCAGCACTGTGCGG 300
 DB 361 AGCAACCTGGGCTGTGCTGGGCAACTCTGCGGCTGTGGCCCGCCAGCACTGTGCGG 420
 QY 301 CACCTGGCGGCAAGCGGCGCGGCGGAGTGTCTTCAAGAACCTATAGTATGACACTTCC 360
 DB 421 CACCTGGCGGCAAGCGGCGCGGCGGAGTGTCTTCAAGAACCTATAGTATGACACTTCC 480
 QY 361 AGCTCTTCAAAAGAGACAGAGGGTAGTGTGCTGCGCGCTCGGAGTCAAGCACTTGTCA 420
 DB 481 AGCTCTTCAAAAGAGACAGAGGGTAGTGTGCTGCGCGCTCGGAGTCAAGCACTTGTCA 540
 QY 421 AATTCTCAGAGGGTCACTGGGAGACAGGCTTCCCTCCCAACCAAGCGGCGGAGT 480
 DB 541 AATTCTCAGAGGGTCACTGGGAGACAGGCTTCCCTCCCAACCAAGCGGCGGAGT 600
 QY 481 CGGGGGCGGGCCAGTGTGTGTCAGACGGGGCGGAGAGGGGGCCCGAGCGGACCCGAG 540
 DB 601 CGGGGGCGGGCCAGTGTGTGTCAGACGGGGCGGAGAGGGGGCCCGAGCGGACCCGAG 660
 QY 541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAAAGTGACTGTGACATCCGCTC 600

Db 661 CAGAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAAAGTGAACCTGTGACATCCGGCTC 720
Qy 601 CGGGTTGAGAGAGTACTGTGAGACATGGGCGACGCTTGGAGCAGGGGTGGCAATCCCGG 660
Db 721 CGGGTTGAGAGAGTACTGTGAGACATGGGCGACGCTTGGAGCAGGGGTGGCAATCCCGG 780
Qy 661 CGGCCCCAGAGCGCTGAGCGCGCAGCTGAGACGTGTTTGGGCGAGCCACCGCAGTGTGCGC 720
Db 781 CGGCCCCAGAGCGCTGAGCGCGCAGCTGAGACGTGTTTGGGCGAGCCACCGCAGTGTGCGC 840
Qy 721 TCAGGAGACTGGGCTGTGTGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGAGC 780
Db 841 TCAGGAGACTGGGCTGTGTGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGAGC 900
Qy 781 GCCCTTGGGGCGACTACCTGAGTGGGCGCCTGCTGACAGGCGCTGGCGGGCGTGTCTCG 840
Db 901 GCCCTTGGGGCGACTACCTGAGTGGGCGCCTGCTGACAGGCGCTGGCGGGCGTGTCTCG 960
Qy 841 ACTGAGGCCCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGTGTCAGTGTGAT 900
Db 961 ACTGAGGCCCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGTGTCAGTGTGAT 1020
Qy 901 GAGGCTGAC 909
Db 1021 GAGGCTGAC 1029

RESULT 9

AAH3283
ID AAH3283 standard; cDNA; 2045 BP.

XX AAH3283;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:339.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KM colorectal carcinoma; ss.

XX Homo sapiens.

OS MO200122920-A2.

PN 05-APR-2001.

XX 28-SEP-2000; 2000MO-US026524.

XX 29-SEP-1999; 99US-0157137P.

PR 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCT INC.

PA Ruben SM, Barash SC, Birze CE, Rosen CA;

XX WPI; 2001-235357/24.

DR P-PSDB; AAG73852.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers.

XX Claim 1; Page 2451; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
cancer-associated nucleic acid molecules (N) and proteins (P), where the
proteins are collectively known as colon cancer antigens. The colon
cancer antigens have cytosolic activity and can be used in gene therapy
and vaccine production. N and P may be used in the prevention, diagnosis
and treatment of diseases associated with inappropriate P expression. For
example, N and P may be used to treat disorders associated with decreased
expression by rectifying mutations or deletions in a patient's genome
that affect the activity of P by expressing inactive proteins or to

CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX

SO Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;

Query Match 99.8%; Score 907.4; DB 4; Length 2045;

Best Local Similarity 99.9%; Pred. No. 4.6e-171;

Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGAGATGAGTGGTGGACTAC 60
Db 121 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGAGATGAGTGGTGGACTAC 180
Qy 61 TACGAGATGCTGTGCTTCAACCGTATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
Db 181 TACGAGATGCTGTGCTTCAACCGTATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 240
Qy 121 GAGCTGAGCTCTTGACCTTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTACCC 180
Db 241 GAGCTGAGCTCTTGACCTTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTACCC 300
Qy 181 CGGGCCCGCAGCGGCTTGAAGCTCTGTGAGAGTGGAGCGCGCGGCACTGAGCGAG 240
Db 301 CGGGCCCGCAGCGGCTTGAAGCTCTGTGAGAGTGGAGCGCGCGGCACTGAGCGAG 360
Qy 241 AGCAACCTGGGCGCTGAGGCGCAACTCTGCGCGCTGCGCGCGCGCAGACCTGCGCG 300
Db 361 AGCAACCTGGGCGCTGAGGCGCAACTCTGCGCGCTGCGCGCGCGCAGACCTGCGCG 420
Qy 301 CACCTGCGCGCAGCGCGCGCGCGCAGTGTCTCAGAACGCTATGAGCACTTCC 360
Db 421 CACCTGCGCGCAGCGCGCGCGCGCAGTGTCTCAGAACGCTATGAGCACTTCC 480
Qy 361 AGCTCTTCAAAAGAGAGAGAGAGGTTAGTCCGCTGCGCGCTGAGAGTTCAGCA 420
Db 481 AGCTCTTCAAAAGAGAGAGAGGTTAGTCCGCTGCGCGCTGAGAGTTCAGCA 540
Qy 421 AATTCTGAGAGGCTCAAGTGGAGAGAGGCTCCCCCAACCAAGCGGCGAGGAGT 480
Db 541 AATTCTGAGAGGCTCAAGTGGAGAGAGGCTCCCCCAACCAAGCGGCGAGGAGT 600
Qy 481 CGGGCCCGCAGTGTGTGTGCGCAGACGCGCGGAGAGGCGCGCAGCGCAG 540
Db 601 CGGGCCCGCAGTGTGTGTGCGCAGACGCGCGGAGAGGCGCGCAGCGCAG 660
Qy 601 CAGCAGTCAAGAGCCCGCAGACCTTCTTGAAGGCAAGTGAACCTGTGACATCCGGTTC 600
Db 661 CAGCAGTCAAGAGCCCGCAGACCTTCTTGAAGGCAAGTGAACCTGTGACATCCGGTTC 720
Qy 601 CGGGTTGAGAGAGTACTGTGAGACATGGGCGACGCTTGGAGCAGGGGTGGCAATCCCGG 660
Db 721 CGGGTTGAGAGAGTACTGTGAGACATGGGCGACGCTTGGAGCAGGGGTGGCAATCCCGG 780
Qy 661 CGGCCCCAGAGCGCTGAGCGCGCAGCTGAGACGTGTTTGGGCGAGCCACCGCAGTGTGCGC 720
Db 781 CGGCCCCAGAGCGCTGAGCGCGCAGCTGAGACGTGTTTGGGCGAGCCACCGCAGTGTGCGC 840
Qy 721 TCAGGAGACTGGGCTGTGTGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGAGC 780
Db 841 TCAGGAGACTGGGCTGTGTGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGAGC 900
Qy 781 GCCCTTGGGGCGACTACCTGAGTGGGCGCCTGCTGACAGGCGCTGGCGGGCGTGTCTCG 840
Db 901 GCCCTTGGGGCGACTACCTGAGTGGGCGCCTGCTGACAGGCGCTGGCGGGCGTGTCTCG 960
Qy 841 ACTGAGGCCCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGTGTCAGTGTGAT 900


```
Db      961 ACTGAGGCTGAGAGGCTGTGAGGCGGAGAGCTGTTCCTGCTGTGATGTGAT 1020
Qy      901 GAGGCTGAC 909
        |||||
Db      1021 GAGGCTGAC 1029

RESULT 10
AAFI8296 standard; DNA; 2044 BP.
XX
AC      AAFI8296;
XX
DT      14-MAR-2001 (first entry)
XX
DE      Lung cancer associated polynucleotide sequence SEQ ID 315.
XX
KW      Human; lung cancer associated protein; neuroprotective; cyrostatic;
KW      cardioactive; immunomodulatory; muscular active; vulnerary;
KW      gastrointestinal; nephrotropic; antiinfective; gynecological;
KW      antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW      proliferative disorder; wound healing; infectious disease; ds.
XX
OS      Homo sapiens.
XX
PN      MO200055180-A2.
XX
PD      21-SEP-2000.
XX
PF      08-MAR-2000; 2000MO-US005918.
XX
PR      12-MAR-1999; 9905-0124270P.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
PA      (ROSE/) ROSEN C A.
XX
PI      Ruben SM;
XX
DR      WPI; 2000-587514/55.
XX      P-PSDB; AAB58420.
XX
PT      Lung cancer associated gene sequences, referred to as lung cancer
PT      antigens, useful for treatment, prevention, and diagnosis of disorders
PT      such as lung cancer.
XX
PS      Claim 1; Page 772-773; 1425pp; English.
XX
CC      Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer
CC      associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC      associated proteins and polynucleotide sequences, their agonists, and
CC      antagonists may have neuroprotective; cyrostatic; cardioactive;
CC      immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC      general; nephrotropic; antiinfective; gynecological; or antibacterial
CC      activity. The invention also includes antibodies specific for the protein
CC      or polynucleotide sequences. The lung cancer associated polynucleotide
CC      sequences may be used for detection of lung cancer, chromosome
CC      identification, as chromosome markers, and for numerous other diagnostic
CC      or research purposes. The proteins may be used to treat disorders such as
CC      neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC      cardiovascular, renal, and proliferative disorders. The proteins may also
CC      be used in the treatment of wounds and infectious diseases.
CC      Polynucleotide sequences AAFI8425 - AAFI8433 and peptide AAB58549 are
CC      used in the course of the invention for the identification and
CC      characterisation of the polynucleotide and protein sequences
XX
SQ      Sequence 2044 BP; 457 A; 620 C; 603 G; 364 T; 0 U; 0 Other;

Query Match      98.5%; Score 895.4; DB 3; Length 2044;
Best Local Similarity 99.8%; Pred. No. 1,1e-168;
Matches 907; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy      1 ATGGCGCTATCCGGGCTGACCCCGGCCCTGCTGTGGAGAGATGATGCTCTGACTAC 60
```

```
Db      121 ATGGCGCTATCCGGGCTGACCCCGGCCCTGCTGTGGAGAGATGATGCTCTGACTAC 180
Qy      61 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGGCGGCACTGACCGAGTGC 120
Db      181 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGGCGGCACTGACCGAGTGC 240
Qy      121 GAGCTGAGGCTCTGACCTTTTCTGTGATGAGGCTCTGAGCGCCCGGAGGCTTAAGC 180
Db      241 GAGCTGAGGCTCTGACCTTTTCTGTGATGAGGCTCTGAGCGCCCGGAGGCTTAAGC 300
Qy      181 CGGGCCCGAGCGGCTTAGAGCTCTGTGAGTGAAGGCTCTGAGCGCCCGGAGGCTTAAGC 240
Db      301 CGGGCCCGAGCGGCTTAGAGCTCTGTGAGTGAAGGCTCTGAGCGCCCGGAGGCTTAAGC 360
Qy      241 AGCAACCTGCGGCTGTGGGCGCAACTCTGCGGCTGAGCCCGCCACGACTGTGAGCG 300
Db      361 AGCAACCTGCGGCTGTGGGCGCAACTCTGCGGCTGAGCCCGCCACGACTGTGAGCG 420
Qy      301 CACTTGCGCGCAGCGGCGCGCCGCAAGTGTCTCCAGAACGCTATAGTACACTCTCC 360
Db      421 CACTTGCGCGCAGCGGCGCGCCGCAAGTGTCTCCAGAACGCTATAGTACACTCTCC 480
Qy      361 AGCTCTTCAAGGAGCAGAGGAGTGTGCTGCGCGCTGCGGCAATGCAAGTTCGCA 420
Db      481 AGCTCTTCAAGGAGCAGAGGAGTGTGCTGCGCGCTGCGGCAATGCAAGTTCGCA 540
Qy      421 AATTCTCAGCAGAGGTGATGAGGAGAGCAGGCTCCCCCAACCAAGCGGACGCGGAGT 480
Db      541 AATTCTCAGCAGAGGTGATGAGGAGAGCAGGCTCCCCCAACCAAGCGGACGCGGAGT 600
Qy      481 CGGGGCGCGGCCAGTGTGTGTCAGAGCGGCGGAGAGGGGCCCAAGCCGACCCCA 540
Db      601 CGGGGCGCGGCCAGTGTGTGTCAGAGCGGCGGAGAGGGGCCCAAGCCGACCCCA 660
Qy      541 CAGCAGTCAAGAGCCCGCCAGACTTCTCTGAAAGGCAAGTGAATCTGTGACATCCGACTC 600
Db      661 CAGCAGTCAAGAGCCCGCCAGACTTCTCTGAAAGGCAAGTGAATCTGTGACATCCGACTC 720
Qy      601 CGGGTTCGAGCAGAGTACTGCGAGCATGAGGCCAGCCTTGAGAGCAGGCGCTGACCCGG 660
Db      721 CGGGTTCGAGCAGAGTACTGCGAGCATGAGGCCAGCCTTGAGAGCAGGCGCTGACCCGG 780
Qy      661 CGGCCCCAGGCGCTGCGCGGCGCAGCTGAGCGTGTGGGAGGCAACCGAGTGTGCGC 720
Db      781 CGGCCCCAGGCGCTGCGCGGCGCAGCTGAGCGTGTGGGAGGCAACCGAGTGTGCGC 840
Qy      721 TCAAGGAGCCTGGGCTGTGTGTGATCATCAAGTTCACAGGCTCTTATGAGAC 780
Db      841 TCAAGGAGCCTGGGCTGTGTGTGATCATCAAGTTCACAGGCTCTTATGAGAC 900
Qy      781 GCCTTTCGAGGCGCAGTACTGAGTGGCGGCTCTGTCAGAGCCCTGCGGGCGTGTCTTG 840
Db      901 GCCTTTCGAGGCGCAGTACTGAGTGGCGGCTCTGTCAGAGCCCTGCGGGCGTGTCTTG 959
Qy      841 ACTGAGGCGCTGTGAGAGGCTGTGGCGCGGAGAGGCTGTGCTGCTGTGATGTGAT 900
Db      960 ACTGAGGCGCTGTGAGAGGCTGTGGCGCGGAGAGGCTGTGCTGCTGTGATGTGAT 1019
Qy      901 GAGGCTGAC 909
        |||||
Db      1020 GAGGCTGAC 1028

RESULT 11
ABL39692 standard; cDNA; 1084 BP.
XX
ID      ABL39692
XX
AC      ABL39692;
XX
DT      10-MAY-2002 (first entry)
XX
DE      Human NS cDNA sequence SEQ ID NO:2.
```


KM antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiac;
 KM anticonvulsant; antidiabetic; tranquilizer; antidepressant; antiepileptic;
 KM gastrointestinal; virucide; antitumor; cerebroprotective; nootropic;
 KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KM rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KM infertility; cardiovascular disease; coagulation disease; hypertension;
 KM ischemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KM gastric ulcer; Alzheimer's disease; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200206315-A2.
 XX
 PD 24-JAN-2002.
 XX
 PF 17-JUL-2001; 2001MO-IL000653.
 XX
 PR 18-JUL-2000; 2000IL-00137345.
 XX
 PR 15-DEC-2000; 2000IL-00140354.
 XX
 PA (COMP-) COMPUGEN LTD.
 PI Mintz L, Freilich S, Bernstein J;
 XX
 DR WPI; 2002-155037/20.
 XX
 PT P-PSDB; ABB06039.
 PT One hundred and twenty eight novel nucleic acid sequences, useful for
 treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 XX
 PS Claim 1; Page 77; 290pp; English.
 XX

CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopapnic, gynaecological, neuroprotective,
 CC antihemmatic, antihypertensive, antiparasitic, ophthalmological, virucide,
 CC vasorelaxant, antiatherosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
 CC antioagulant, antifibrinolytic, hypotension, antiasthmatic, cardiac,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilizer, antitumor,
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC nootropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and anticancer therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcer, stroke,
 CC Alzheimer's disease and as a contraceptive
 XX

XX Sequence 1106 BP; 182 A; 340 C; 385 G; 189 T; 0 U; 10 Other;

Query Match 90.8%; Score 825.8; DB 6; Length 1106;
 Best Local Similarity 98.0%; Pred. No. 7.3e-155;
 Matches 891; Conservative 7; Mismatches 4; Indels 7; Gaps 6;

QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGAGATGCTCTGACTAC 60
 DB 106 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGAGATGCTCTGACTAC 165
 QY 61 TAGGGATGCTGTGCTTACCGTATTTGAGGTGTGGCGGCACTGACCGAGTGC 120
 DB 166 TAGGGATGCTGTGCTTACCGTATTTGAGGTGTGGCGGCACTGACCGAGTGC 225
 QY 121 GAGGTGAGCTCCGCTTCTGCTGGATGAGGCTCTGAGCGCGCGAGGCTTAGCC 180
 DB 226 GAGGTGAGCTCCGCTTCTGCTGGATGAGGCTCTGAGCGCGCGAGGCTTAGCC 284

QY 181 CGGGCCCGACCGGCTTAAAGCTCTGCTGAGACTGAGCGCCGCGGCACTGCGCGAG 240
 DB 285 CGGGCCCGACCGGCTTAAAGCTCTGCTGAGACTGAGCGCGCGGCACTGCGCGAG 344
 QY 241 AGCAACCTGGGTGTGGGGGCAACCTGCGGCTGCTGGCCCGGCACTGCTGGCCG 300
 DB 345 AGCAACCTGGGTGTGGGGGCAACCTGCGGCTGCTGGCCCGGCACTGCTGGCCG 404
 QY 301 CACCTGCGCGACGAGCGCGCCCGGCACTGCTGCAAGACGCTATGACTAGGCACTCC 360
 DB 405 CACCTGCGCGACGAGCGCGCCCGGCACTGCTGCAAGACGCTATGACTAGGCACTCC 462
 QY 361 AGCTTTCAAAAGAGACAGAGGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCA 420
 DB 463 AGCTTTCAAAAGAGAGACAGAGGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCA 522
 QY 421 AATTCTCAGAGGAGTGTGAGGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 480
 DB 523 AATTCTCAGAGGAGTGTGAGGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 581
 QY 481 CGGGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 540
 DB 582 CGGGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 640
 QY 541 CAGCAGTCAAGAGCGCGCAGACCTTCTCTGAAAGCAAGTGAATCTGTGATCCGCTC 600
 DB 641 CAGCAGTCAAGAGCGCGCAGACCTTCTCTGAAAGCAAGTGAATCTGTGATCCGCTC 700
 QY 601 CGGGTTGAGAGAGTGTGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 DB 701 CGGGTTGAGAGAGTGTGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 760
 QY 661 CGGGCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 720
 DB 761 CGG-CCGAGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 819
 QY 721 TCAAGGACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 780
 DB 820 TCAAGGACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 879
 QY 781 GCCTTGTGGGGGAGTACCTGAGTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 DB 880 GCCTTGTGGGGGAGTACCTGAGTGTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 939
 QY 841 ACTGAGGCGCTGCGAGAGGCTGTGGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 940 ACTGAGGCGCTGCGAGAGGCTGTGGCGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 998
 QY 901 GAGGCTGAC 909
 DB 999 GAGGCTGAC 1007

RESULT 13

AA562603 standard; cDNA; 1966 BP.

AA562603;

14-FEB-2002 (first entry)

cDNA sequence #390 encoding novel human secreted protein.

Human secreted protein; hyperproliferative disorder; autoimmune disorder;

immune deficiency disorder; blood disorder; inflammatory disorder;

infectious disorder; gene therapy; antimicrobial; hepatocytic;

immunosuppressive; antineumatic; ss.

Homo sapiens.
 W0200177291-A2.

PD 18-OCT-2001.
 XX 29-MAR-2001; 2001WO-US010485.
 XX 06-APR-2000; 2000US-0195604P.
 XX (GBMY) GENETICS INST INC.
 PA
 PI Wong GG, Clark HF, Fehchel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;
 XX WPI; 2002-010900/01.
 XX
 XX New polynucleotides encoding secreted proteins useful for treating e.g.
 PT asthma, HIV and Crohn's disease.
 DR
 XX
 XX
 PS Claim 1; Page 280; 39pp; English.
 CC The present invention relates to the isolation of novel cDNA sequences
 CC which encode human secreted proteins. The cDNA sequences have been
 CC derived from a variety of human tissues. The invention also provides a
 CC method for producing proteins from these polynucleotide sequences. The
 CC proteins are useful for identifying compounds that modulate their
 CC activity and production, and the cell is also useful for identifying
 CC compounds that modulate expression of the polynucleotide sequences
 CC encoding the secreted proteins. The sequences of the invention are useful
 CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
 CC immune deficiency disorders (e.g. severe combined immunodeficiency
 CC (SCID), autoimmune disorders (e.g. multiple sclerosis), blood disorders
 CC (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis) and
 CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
 CC the invention are also useful in gene therapy. AAS62214-AAS62838.
 CC represent the cDNA sequences of the invention that encode for novel human
 CC secreted proteins
 XX
 SQ Sequence 1966 BP; 356 A; 618 C; 614 G; 378 T; 0 U; 0 Other;
 Query Match 87.9%; Score 799.4; DB 6; Length 1966;
 Best Local Similarity 90.2%; Pred. No. 1.3e-149;
 Matches 908; Conservative 0; Mismatches 1; Indels 98; Gaps 1;
 QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGATGAGTGGCTGACTAC 60
 DB 50 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGATGAGTGGCTGACTAC 109
 QY 61 TACGGGATGCTGTGCTTACCGATGTGTGAGGTGGGCGGCACTGACCGAGTGC 120
 DB 110 TACGGGATGCTGTGCTTACCGATGTGTGAGGTGGGCGGCACTGACCGAGTGC 169
 QY 121 GAGCTGAGAGCTCTGGCTTTCTCTGTGATGAGGCTCTGGCGCGCGAGGCTTACC 180
 DB 170 GAGCTGAGAGCTCTGGCTTTCTCTGTGATGAGGCTCTGGCGCGCGAGGCTTACC 229
 QY 181 CGGGCCCGCAGCGGCTTGAAGCTCTGTGTGAGCTGAGCGCGCGGCACTGCGCGAG 240
 DB 230 CGGGCCCGCAGCGGCTTGAAGCTCTGTGTGAGCTGAGCGCGCGGCACTGCGCGAG 289
 QY 241 AGCAACTGGCGGCTGTGGGCACTCTGTGGCGTGTGGCGCGCGCAACTGTCTGCGG 300
 DB 290 AGCAACTGGCGGCTGTGGGCACTCTGTGGCGTGTGGCGCGCGCAACTGTCTGCGG 349
 QY 301 CACCTGCGCGCAGCGGCGCGGCACTGTCTCAGAACTTATAGTATGAGCACTTCC 360
 DB 350 CACCTGCGCGCAGCGGCGCGGCACTGTCTCAGAACTTATAGTATGAGCACTTCC 409
 QY 361 AGCTTTCAAAGAGAGAGAGAGTGTGCGGTGCGGTGCGGAGTCAAGCACTTTCGA 420
 DB 410 AGCTTTCAAAGAGAGAGAGAGTGTGCGGTGCGGTGCGGAGTCAAGCACTTTCGA 469
 QY 421 AATTCTGAGCGAGGTGAGTGGAGAGAGAGTCCCCCAACAGCGGCAAGCGGAGT 480
 DB 470 AATTCTGAGCGAGGTGAGTGGAGAGAGAGTCCCCCAACAGCGGCAAGCGGAGT 529

QY 481 CGGGCCCGCCAGTGTGTGTCAGACGCGCGGAGAGAGGCCCCAGCCGACCCAG 540
 DB 530 CGGGCCCGCCAGTGTGTGTCAGACGCGCGGAGAGAGGCCCCAGCCGACCCAG 589
 QY 541 CAGAGTCAGAGCCCGCAGACCTTCTGTAAGCAAGGACTG----- 587
 DB 590 CAGAGTCAGAGCCCGCAGACCTTCTGTAAGCAAGGACTGCTGTACAAAGA 649
 QY 588 ----- 587
 DB 650 AGCAGGATGCCAGATGTGCTTGTGAGGAACTCCGGAAGCTTCATTCATGTGGA 709
 QY 588 -----TGACATCCGCTTCGGGTTTCAGAGAGTATGTCG 622
 DB 710 GGCAGAGGAGAGAGAGCTGTGTCAATGATCCGGGCTCCGGGTTTCAGAGAGTATGTCG 769
 QY 623 AGCATGGGCGCAGCTTTGAGACAGGCGCTGCAATCCCGCGCGCCAGAGCGCTGCGCGC 682
 DB 770 AGCATGGGCGCAGCTTTGAGACAGGCGCTGCAATCCCGCGCGCCAGAGCGCTGCGCGC 829
 QY 683 AGCTGACGATGTTGGGAGAGCCAGCGAGTGTGCGCTCAAGGAGCTTGGGCTGTGG 742
 DB 830 AGCTGACGATGTTGGGAGAGCCAGCGAGTGTGCGCTCAAGGAGCTTGGGCTGTGG 889
 QY 743 TTTGTGACATCAAGTCTTCAAGCTCTTATCTGAGCGCTTTCGAGGAGCACTACCTGA 802
 DB 890 TTTGTGACATCAAGTCTTCAAGCTCTTATCTGAGCGCTTTCGAGGAGCACTACCTGA 949
 QY 803 GTGGCGCCCTGTGTCAGAGCCCTGTGGGCGGTGTCTGACTGAGAGCCCTGGAAGGCTG 862
 DB 950 GTGGCGCCCTGTGTCAGAGCCCTGTGGGCGGTGTCTGACTGAGAGCCCTGGAAGGCTG 1009
 QY 863 TGAGCGGAGAGGCTTTCGAGCTGTGAGTGTGAGTGAAGGCTGAC 909
 DB 1010 TGAGCGGAGAGGCTTTCGAGCTGTGAGTGTGAGTGAAGGCTGAC 1056
 RESULT 14
 AAH99646
 ID AAH99646 standard; cDNA; 1570 BP.
 XX
 XX AAH99646;
 XX
 XX 16-OCT-2001 (first entry)
 DE Human protein encoding cDNA sequence SEQ ID NO:481.
 XX
 XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiac; central nervous system; vitruide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KW antiaggregant; haemostatic; vulnery; antidiabetic; osteopathic; eczema;
 KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antitense therapy; vaccine; inflammation;
 KW anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; hematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.
 XX
 XX Homo sapiens.
 OS
 XX MO200153455-A2.
 XX
 XX 26-JUL-2001.
 XX
 XX 22-DEC-2000; 2000MO-US035017.
 XX
 XX 23-DEC-1999; 99US-00471275.
 XX
 XX 21-JAN-2000; 2000US-00488725.
 PR

PR 25-APR-2000; 2000US-00552317.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI: 2001-457603/49.
 XX P-PSDB; AAM25705.
 DR
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 XX
 PS Claim 1; Page 552; 1217pp; English.
 XX
 XX AAH9166 to AAH9904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antiinflammatory; antineurotic;
 CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiatic;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; valvularty;
 CC antitumor; osteoplastic; dermatological; antiallergic; antilethemic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX
 XX Sequence 1570 BP; 311 A; 488 C; 466 G; 305 T; 0 U; 0 Other;
 SQ
 Query Match 63.9%; Score 581.2; DB 4; Length 1570;
 Best Local Similarity 96.3%; Pred. No. 3.1e-106;
 Matches 595; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 292 CMCCTCCGACCTGCGGCGCAAGCGCGCCGCACTGCTCCAGAAAGCTTATGCTAT 351
 DB 40 CGGCGCGCGCCAGAGTCATCGAGCGCAAAATCTGTGCTCCAGAAACCTTATGCTAT 99
 QY 352 GGCACCTCCAGCTCTTCAAGAGAGACAGAGGTAGCTGCGCGCTCGGAGTCAAGC 411
 DB 100 GGCACCTCCAGCTCTTCAAGAGAGACAGAGGTAGCTGCGCGCTCGGAGTCAAGC 159
 QY 412 AGTTCTGCAAAATTTCTCAGCAGGGTCACTGAGAGACAGGCTTCCCCCAACCAAGCGGAG 471
 DB 160 AGTTCTGCAAAATTTCTCAGCAGGGTCACTGAGAGACAGGCTTCCCCCAACCAAGCGGAG 219
 QY 472 CGGCGAGTGGGGCGGCGCGGAGTGTGTGCGAGAGGGGGGAGAGGGGGCGGAGCC 511
 DB 220 CGGCGAGTGGGGCGGCGCGGAGTGTGTGCGAGAGGGGGGAGAGGGGGCGGAGCC 279
 QY 532 GCACCCGAGAGAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAAAGTACTGTGAC 591
 DB 280 GCACCCGAGAGAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAAAGTACTGTGAC 339
 QY 592 ATCCGGCTCCGGGTTCCAGAGAGTACTGCGAGATGGGCGAGCTTGAAGCAGGGCGTG 651
 DB 340 ATCCGGCTCCGGGTTCCAGAGAGTACTGCGAGATGGGCGAGCTTGAAGCAGGGCGTG 399
 QY 652 GCATCCCGGGGGCGGCGGCGGCTGTGAGAGTGTGTTGGGAGGCGCACCGCA 711
 DB 400 GCATCCCGGGGGCGGCGGCGGCTGTGAGAGTGTGTTGGGAGGCGCACCGCA 459
 QY 712 GTCTGCGCTCAAGAGAGTGTGTTGTGATCAATCAAGTTCTCAGAGCTTCC 771
 DB 460 GTCTGCGCTCAAGAGAGTGTGTTGTGATCAATCAAGTTCTCAGAGCTTCC 519

QY 772 TATCTGAGCGCTTCTGCGGCGCACTACTGAGTGGGCGCCCTGCTGCAGAGCGCTCGGGGC 831
 DB 520 TATCTGAGCGCTTCTGCGGCGCACTACTGAGTGGGCGCCCTGCTGCAGAGCGCTCGGGGC 579
 QY 832 GTGTTCTGACTAGAGCGCTTGCAGAGAGGCTGTGGGCGGAGAGGCTGTGCTGTGCTGTC 891
 DB 580 GTGTTCTGACTAGAGCGCTTGCAGAGAGGCTGTGGGCGGAGAGGCTGTGCTGTGCTGTC 639
 QY 892 AGTGTGATGAGGCTGAC 909
 DB 640 AGTGTGATGAGGCTGAC 657
 RESULT 15
 AA158961
 ID AA158961 standard; cDNA; 1067 BP.
 XX
 AC AA158961;
 XX
 XX 22-OCT-2001 (first entry)
 DE
 XX Human polynucleotide SEQ ID NO 1164.
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153112-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US034263.
 XX
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM39805.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Claim 1; SEQ ID NO 1164; 10078pp; English.
 XX
 XX The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM38642-AA42213) with nootropic,
 CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX
SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;

Query Match 59.5%; Score 541; DB 4; Length 1067;
Best Local Similarity 95.8%; Pred. No. 3e-98;
Matches 572; Conservative 0; Mismatches 10; Indels 15; Gaps 1;

```
QY 1 ATGGCGCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 60
   |||||
Db 151 ATGGCGCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 210
   |||||

QY 61 TACGGGATGCTGTGCTTCAACCGTATGTCAGAGTGTGGGCGGGCAACTGACCGAGTGC 120
   |||||
Db 211 TACGGGATGCTGTGCTTCAACCGTATGTCAGAGTGTGGGCGGGCAACTGACCGAGTGC 270
   |||||

QY 121 GAGCTGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAGCC 180
   |||||
Db 271 GAGCTGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAGCC 330
   |||||

QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCCGCGGCACTGCGCGAG 240
   |||||
Db 331 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCCGCGGCACTGCGCGAG 390
   |||||

QY 241 AGCAACTTGGGGCTGTGGGGCAACTCTGTGCGCTGTGGCGCCGACAGACTGTGCG 300
   |||||
Db 391 AGCAACTTGGGGCTGTGGGGCAACTCTGTGCGCTGTGGCGCCGACAGACTGTGCG 450
   |||||

QY 301 CACCTGCGCGCAGCGGCGCGGCAAGTGTCTCCAGAACGCTATAGCTATGGCACTCC 360
   |||||
Db 451 CACCTGCGCGCAGCGGCGCGGCAAGTGTCTCCAGAACGCTATAGCTATGGCACTCC 510
   |||||

QY 361 AGCTCTTCAAAGAGAGAGAGGTAAGTCCGTCGCGTCCGAGTCAAGCAGTTTGCA 420
   |||||
Db 511 AGCTCTTCAAAGAGAGAGAGGTAAGTCCGTCGCGTCCGAGTCAAGCAGTTTGCA 570
   |||||

QY 421 AATTCTCAGCAGGATCAGTGGAGACAGGCTCCCCCAACCAAGCGGCAAGCGGAGT 480
   |||||
Db 571 AATTCTCAGCAG-----GGCTCCCCCAACCAAGCGGCAAGCGGAGT 615
   |||||

QY 481 CGGGGCGGCGCAGTGTGTGCCAGACGGCGGCGAGAGGGGCCCAAGCCGCAAGCCGAG 540
   |||||
Db 616 CGGGGCGGCGCAGTGTGTGCCAGACGGCGGCGAGAGGGGCCCAAGCCGCAAGCCGAG 675
   |||||

QY 541 CAGCAGTCAGAGCCCGCAGACTTCTCTGAAAGCAAAAGTGAACCTGTGACATCCGG 597
   |||||
Db 676 CAGCAGTCAGAGCCCGCAGACTTCTCTGAAAGCAAAAGTGAACCAACCAACTG 732
   |||||
```

Search completed: March 23, 2005, 09:52:55
Job time : 535.452 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 09:03:21 ; Search time 3257.36 Seconds
(without alignments)
10622.233 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atgagcgctatccggatcgac.....tcagtgatgagagctgac

Scoring table: IDENTITY NUC

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907.4	99.8	1053	3	CR604307 full-length
2	907.4	99.8	1860	3	CR593642 full-length
3	907.4	99.8	1866	3	CR619301 full-length
4	907.4	99.8	1894	3	CR625070 full-length
5	847.2	93.2	1046	5	CR625070 full-length
6	822.8	90.5	1021	5	CR625070 full-length
7	802	88.2	1091	5	CR625070 full-length
8	773.6	85.1	951	5	CR625070 full-length
9	760.4	83.7	800	4	CR625070 full-length
10	728.4	81.2	837	4	CR625070 full-length
11	727.4	80.0	1089	5	CR625070 full-length
12	711.2	78.2	820	5	CR625070 full-length
13	704.2	77.5	1090	5	CR625070 full-length
14	680.8	74.9	872	4	CR625070 full-length
15	678.8	74.7	874	1	CR625070 full-length
16	653	71.8	924	5	CR625070 full-length
17	639.6	70.4	931	5	CR625070 full-length
18	624.2	68.7	718	4	CR625070 full-length
19	607.8	66.9	1042	5	CR625070 full-length
20	605	66.6	926	4	CR625070 full-length
21	596.4	65.6	1641	3	CR625070 full-length
22	593.2	65.3	1025	6	CR625070 full-length
23	588.8	64.8	4598	3	CR625070 full-length
24	586.8	64.6	878	4	CR625070 full-length

25	582.8	64.1	905	4	BI910528
26	582.8	64.1	927	4	BI910528
27	582.8	64.1	947	4	BI910528
28	582.8	64.1	1566	3	BI910528
29	582.4	64.1	634	4	BI910528
30	582.4	64.1	1131	2	BI910528
31	582.4	64.1	1510	3	BI910528
32	582	64.0	1062	5	BI910528
33	581.4	64.0	655	7	BI910528
34	581.2	63.9	1051	5	BI910528
35	581.2	63.9	1161	5	BI910528
36	580.8	63.9	644	7	BI910528
37	579.6	63.8	678	7	BI910528
38	579.6	63.8	783	5	BI910528
39	577.4	63.5	826	6	BI910528
40	574	63.1	938	5	BI910528
41	572.4	63.0	671	7	BI910528
42	570.6	62.8	780	6	BI910528
43	570.4	62.8	931	5	BI910528
44	569.4	62.6	993	4	BI910528
45	562.8	61.9	1017	3	BI910528

ALIGNMENTS

RESULT 1
LOCUS CR604307 1053 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DL006YU09 of B cells (Ramos cell line)
ACCESSION CR604307
VERSION CR604307.1 GI:50485114
KEYWORDS HTC; CNSLT; CDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Li, W.B., Gruber, C., Jessee, J. and Polayres, D.
1 (bases 1 to 1053)
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@life.rockefeller.edu
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue
2 (bases 1 to 1053)
REFERENCE
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage ;
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1053
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL006YU09"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 99.8%; Score 907.4; DB 3; Length 1053;
Best Local Similarity 99.8%; Pred. No. 2.9e-189;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Ct 1 ATGCGCTATCCGGGATCGACCCCGCGCTGCGGAGAGATGAGTGGCTGACTAC 60
Db 99 ATGCGCTATCCGGGATCGACCCCGCGCTGCGGAGAGATGAGTGGCTGACTAC 158

```

QY 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGCTGCGCGGCACTGACCGAGTGC 120
DB 159 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGCTGCGCGGCACTGACCGAGTGC 218
QY 121 GACCTGAGAGTCCGCGCTTCTGCTGAGTGAAGGCTCTGCGCGCGCGGAGGCTTAGCC 180
DB 219 GACCTGAGAGTCCGCGCTTCTGCTGAGTGAAGGCTCTGCGCGCGCGGAGGCTTAGCC 278
QY 181 CGGCGCCGCAAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGCGGCACTGAGCGGAG 240
DB 279 CGGCGCCGCAAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGCGGCACTGAGCGGAG 338
QY 241 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCTGCGCGCGGCACTGCGGCG 300
DB 339 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCTGCGCGCGGCACTGCGGCG 398
QY 301 CACCTGCGGCAAGCGGCGCGGCTGCTGCTGAGCTGAGCTGAGCTGAGCGGAGTGC 360
DB 399 CACCTGCGGCAAGCGGCGCGGCTGCTGCTGAGCTGAGCTGAGCTGAGCGGAGTGC 458
QY 361 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 420
DB 459 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 518
QY 421 AATTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 480
DB 519 AATTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 578
QY 481 CGGCGCGCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 579 CGGCGCGCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 638
QY 541 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 600
DB 639 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 698
QY 601 CGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 660
DB 699 CGGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 758
QY 661 CGGCGCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 759 CGGCGCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 818
QY 721 TCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 780
DB 819 TCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 878
QY 781 GCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 840
DB 879 GCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 938
QY 841 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 900
DB 939 ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 998
QY 901 GAGAGTCAAG 909
DB 999 GAGAGTCAAG 1007

```

```

RESULT 2
CR593642
LOCUS full-length cDNA clone CS0DK010YK16 of Hela cells Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR593642
VERSION CR593642.1 GI:50474449
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS 1 (bases 1 to 1860)
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
REMARK Full-length cDNA libraries and normalization
Unpublished
Contact: Peng Liang Email: liang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1860)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submision
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
COMMENT
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010YK16"
/cisue_type="Hela cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 99.8%; Score 907.4; DB 3; Length 1860;
Best Local Similarity 99.9%; Pred. No. 3e-189; 1; Indels 0; Gaps 0;
Matches 908; Conservative 0; Mismatches 1;
QY 1 ATGCGCTATCCCGGCTGAGACCCCGGCGCGGCTGAGAGAGAGATGCTGAGCTAC 60
DB 71 ATGCGCTATCCCGGCTGAGACCCCGGCGCGGCTGAGAGAGAGATGCTGAGCTAC 130
QY 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGCTGCGCGGCACTGACCGAGTGC 120
DB 131 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGCTGCGCGGCACTGACCGAGTGC 190
QY 121 GACCTGAGAGTCCGCGCTTCTGCTGAGTGAAGGCTCTGCGCGCGGCACTGAGGCTTAGCC 180
DB 191 GACCTGAGAGTCCGCGCTTCTGCTGAGTGAAGGCTCTGCGCGCGGCACTGAGGCTTAGCC 250
QY 191 GACCTGAGAGTCCGCGCTTCTGCTGAGTGAAGGCTCTGCGCGCGGCACTGAGGCTTAGCC 250
DB 181 CGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 240
DB 251 CGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 310
QY 241 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCTGCGCGGCACTGCGGCGG 300
DB 311 AGCAACCTGCGGCTGCTGCGGCACTCTGCGCGGCTGCGCGGCACTGCGGCGG 370
QY 301 CACCTGCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 360
DB 371 CACCTGCGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGT 430
QY 361 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 420
DB 431 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 490
QY 421 AATTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 480
DB 491 AATTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 550
QY 481 CGGCGCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 551 CGGCGCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 610
QY 541 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 600
DB 611 CAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 670

```


QY 601 CGGGTTGAGCAGAGTACTGCGAGCATGGGCGAGCCTTGGAGCAGGGCGTGGCATCCCGG 660
DB 671 CGGGTTGAGCAGAGTACTGCGAGCATGGGCGAGCCTTGGAGCAGGGCGTGGCATCCCGG 730
QY 661 CGGGCCGAGGCGCTGGGGCGGCGAGCTGGAAGTGTGGGCGAGCGACCCGAGTCTGGCG 720
DB 731 CGGGCCGAGGCGCTGGGGCGGCGAGCTGGAAGTGTGGGCGAGCGACCCGAGTCTGGCG 790
QY 721 TCAAGGAGCCTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 780
DB 791 TCAAGGAGCCTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 850
QY 781 GCCTTGGGGCGAGCTACTAGTGGCGCCCTGTGTGACAGGCGCTCGGGGCGTGTCTCG 840
DB 851 GCCTTGGGGCGAGCTACTAGTGGCGCCCTGTGTGACAGGCGCTCGGGGCGTGTCTCG 910
QY 841 ACTGAGGCGCTGGCGAGGCTGTGGGCGGAGGCTGTCCCTGCTGCTGATGTTGGAT 900
DB 911 ACTGAGGCGCTGGCGAGGCTGTGGGCGGAGGCTGTCCCTGCTGCTGATGTTGGAT 970
QY 901 GAGGCTGAC 909
DB 971 GAGGCTGAC 979

RESULT 3
CR619301 1866 bp mRNA linear HTC 21-JUL-2004
LOCUS 25-normalized of Homo sapiens (human)
DEFINITION
ACCESSION CR619301.1 GI:50500108
VERSION HTC; CNSLT CDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1866)
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue
2 (bases 1 to 1866)
Genoscope.
REFERENCE Direct Submission
AUTHORS Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1. 1866
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC015YJ24"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 99.8%; Score 907.4; DB 3; Length 1866;
Best Local Similarity 99.9%; Pred. No. 3e-189;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGGTATCCGGGTGACACCGCGCCCGCTGGCGAGGAGATGATGCTGCTGACTAC 60
DB 76 ATGGCGGTATCCGGGTGACACCGCGCCCGCTGGCGAGGAGATGATGCTGCTGACTAC 135

QY 61 TACGGAGTGTCTGTGCTTACACCGTATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 120
DB 136 TACGGAGTGTCTGTGCTTACACCGTATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 195
QY 121 GAGCTGAGAGCTCTGGGCTTTTCTGCTGATGAGGCTCTGGCGCGCGGAGGCTTGGCC 180
DB 196 GAGCTGAGAGCTCTGGGCTTTTCTGCTGATGAGGCTCTGGCGCGCGGAGGCTTGGCC 255
QY 181 CGGGCCGAGCGGCTTGAAGCTTCTGCTGAGAGTGAAGCGCGCGGAGTGTGGCGAG 240
DB 256 CGGGCCGAGCGGCTTGAAGCTTCTGCTGAGAGTGAAGCGCGCGGAGTGTGGCGAG 315
QY 241 AGCAACCTGGGCTGTGGGCGAACTCTGTGGCGGTGTGCGCCCGCAAGACTGTCCG 300
DB 316 AGCAACCTGGGCTGTGGGCGAACTCTGTGGCGGTGTGCGCCCGCAAGACTGTCCG 375
QY 301 CACCTGGCGGCGAAGCGGCGCGGCGAGTGTCTTCCAAAGCTTATGGAACCTCC 360
DB 376 CACCTGGCGGCGAAGCGGCGCGGCGAGTGTCTTCCAAAGCTTATGGAACCTCC 435
QY 361 AGCTCTTCAAGAAGAGCAGAGAGGTAGCTGCGCGTGTGCGGCGAGTCAAGAGTTCTGCA 420
DB 436 AGCTCTTCAAGAAGAGCAGAGAGGTAGCTGCGCGTGTGCGGCGAGTCAAGAGTTCTGCA 495
QY 421 AATTCTCAGCAGAGGTGAGTGGAGAGCAGGCTCCCCCAACCAAGCGGCGAGGAGT 480
DB 496 AATTCTCAGCAGAGGTGAGTGGAGAGCAGGCTCCCCCAACCAAGCGGCGAGGAGT 555
QY 481 CGGGGCGGCGGCAAGTGTGTGCTGAGACGCGCGGAGAGGAGGCGCCGAGCCCGCAG 540
DB 556 CGGGGCGGCGGCAAGTGTGTGCTGAGACGCGCGGAGAGGAGGCGCCGAGCCCGCAG 615
QY 541 CAGCAGTCAAGAGCGCGGCGAGACCTTCTCTGAAGGCGAAAGTGAAGTGTGACATCCGCTC 600
DB 616 CAGCAGTCAAGAGCGCGGCGAGACCTTCTCTGAAGGCGAAAGTGAAGTGTGACATCCGCTC 675
QY 601 CGGGTTGAGCAGAGTACTGCGAGCATGGGCGAGCCTTGGAGCAGGGCGTGGCATCCCGG 660
DB 676 CGGGTTGAGCAGAGTACTGCGAGCATGGGCGAGCCTTGGAGCAGGGCGTGGCATCCCGG 735
QY 661 CGGGCCGAGGCGCTGGGGCGGCGAGCTGGAAGTGTGGGCGAGCGACCCGAGTCTGGCG 720
DB 736 CGGGCCGAGGCGCTGGGGCGGCGAGCTGGAAGTGTGGGCGAGCGACCCGAGTCTGGCG 795
QY 721 TCAAGGAGCCTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 780
DB 796 TCAAGGAGCCTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 855
QY 781 GCCTTGGGGCGAGCTACTAGTGGCGCCCTGTGTGACAGGCGCTCGGGGCGTGTCTCG 840
DB 856 GCCTTGGGGCGAGCTACTAGTGGCGCCCTGTGTGACAGGCGCTCGGGGCGTGTCTCG 915
QY 841 ACTGAGGCGCTGGCGAGGCTGTGGGCGGAGGCTGTGGCTGCTGCTGCTGATGTTGGAT 900
DB 916 ACTGAGGCGCTGGCGAGGCTGTGGGCGGAGGCTGTGGCTGCTGCTGATGTTGGAT 975
QY 901 GAGGCTGAC 909
DB 976 GAGGCTGAC 984

RESULT 4
CR625070 1894 bp mRNA linear HTC 21-JUL-2004
LOCUS 25-normalized of Homo sapiens (human)
DEFINITION
ACCESSION CR625070.1 GI:50505877
VERSION HTC; CNSLT CDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

REFERENCE 1 (bases 1 to 1894)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1894)
REFERENCE 2 (bases 1 to 1894)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
Source Location/Qualifiers
1..1894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00C001YC18"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 99.8%; Score 907.4; DB 3; Length 1894;
Best Local Similarity 99.9%; Pred. No. 3e-189;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCCCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATAGTCTGAGACTAC 60
DB 89 ATGGCCCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATAGTCTGAGACTAC 148
QY 61 TAGGGATGCTGCTGCTTCAACCGTATGTTGAGATGCTGGGCGGCACTGACCGAGTGC 120
DB 149 TAGGGATGCTGCTGCTTCAACCGTATGTTGAGATGCTGGGCGGCACTGACCGAGTGC 208
QY 121 GAGCTGAGAGCTCTGGGCTTCTGCTGATGATGAGCTCTGGGCGGCGGAGGCTTAGCC 180
DB 209 GAGCTGAGAGCTCTGGGCTTCTGCTGATGATGAGCTCTGGGCGGCGGAGGCTTAGCC 268
QY 181 CGGCGCCGAGCGGCGCTTGAAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGCAGG 240
DB 269 CGGCGCCGAGCGGCGCTTGAAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGCAGG 328
QY 241 AGCAACCTGGGCTGCTGGGGCAATCTCTGCTGCTGGCGCCGACCACTGCTGCCG 300
DB 329 AGCAACCTGGGCTGCTGGGGCAATCTCTGCTGCTGGCGCCGACCACTGCTGCCG 368
QY 301 CACCTGGCGGAGCGGCGCGGCGAGTCTTCAGAAAGCTATAGCTATGACACTTCC 360
DB 389 CACCTGGCGGAGCGGCGCGGCGAGTCTTCAGAAAGCTATAGCTATGACACTTCC 448
QY 361 AGCTTTCAAAAGAGACAGAGGAGTCTGCGCTGCGCGTGGGAGTCAAGCACTTTCGA 420
DB 449 AGCTTTCAAAAGAGACAGAGGAGTCTGCGCTGCGCGTGGGAGTCAAGCACTTTCGA 508
QY 421 AATTTCAGAGAGGCTCAGTGGAGACAGGCTCCCCCAACAAAGCGGCGGAGT 480
DB 509 AATTTCAGAGAGGCTCAGTGGAGACAGGCTCCCCCAACAAAGCGGCGGAGT 568
QY 481 CGGCGCGGCGGCGAGTGGTGGCGAGCGGCGGAGAGAGGCGCCAGCGGAGCCGAG 540
DB 569 CGGCGCGGCGGCGAGTGGTGGCGAGCGGCGGAGAGAGGCGCCAGCGGAGCCGAG 628
QY 541 CAGCAGTCAGAGCCCGGAGAGCTTCTCTGAAGCAAAAGTGAAGCTGTGATCCGCTC 600
DB 629 CAGCAGTCAGAGCCCGGAGAGCTTCTCTGAAGCAAAAGTGAAGCTGTGATCCGCTC 688
QY 601 CGGCTTCGAGCAGAGTCTCGAGCATGGGCGAGCTTGGAGAGAGGCGTGGCATCCGG 660

```

```

DB 689 CGGCTTCGAGCAGAGTCTCGAGCATGGGCGAGCTTGGAGAGAGGCGTGCATCCGG 748
QY 661 CGGCGCCAGCGCTGGCGGCGGAGCTGAGCTGATGCTTTTGGCGAGGCCACCGCAGTGC 720
DB 749 CGGCGCCAGCGCTGGCGGCGGAGCTGAGCTGATGCTTTTGGCGAGGCCACCGCAGTGC 808
QY 721 TCAGAGGAGCTTGGCTCTGCTGTTTGTGACATCAAGTCTCAGAGCTCTTATCTGAGC 780
DB 809 TCAGAGGAGCTTGGCTCTGCTGTTTGTGACATCAAGTCTCAGAGCTCTTATCTGAGC 868
QY 781 GCCTTCGGGGGAGCTACCTGAGTGGCGGCTGCTGTCAGAGCCCTGGCGGGCGTGTCTG 840
DB 869 GCCTTCGGGGGAGCTACCTGAGTGGCGGCTGCTGTCAGAGCCCTGGCGGGCGTGTCTG 928
QY 841 ACTGAGCCCTGCGAGAGGCTGTGGCGGAGAGCTGTTCGCTGCTGCTGATGATGAT 900
DB 929 ACTGAGCCCTGCGAGAGGCTGTGGCGGAGAGGCTGTTCGCTGCTGCTGATGATGAT 988
QY 901 GAGGCTGAC 909
DB 989 GAGGCTGAC 997

```

```

RESULT 5
BX342837 1046 bp mRNA linear EST 07-APR-2004
LOCUS BX342837 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL006YL09 5-PRIME, mRNA sequence.
ACCESSION BX342837
VERSION BX342837.2 GI:46266793
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1046)
Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 1, 2003 this sequence version replaced gi:30313128.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DL006CP05QPlc=9074.r.

```

```

FEATURES
Source Location/Qualifiers
1..1046
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL006YL09"
/cell_line="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

ORIGIN
Query Match 93.2%; Score 847.2; DB 5; Length 1046;
Best Local Similarity 98.1%; Pred. No. 5.2e-176;
Matches 888; Conservative 9; Mismatches 3; Indels 5; Gaps 4;

```

```

QY      1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGATGAGTGCCTGGACTAC 60
Db      99 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGATGAGTGCCTGGACTAC 158
QY      61 TACGGGATGCTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 120
Db      159 TACGGGATGCTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 218
QY      121 GAGCTGAGACTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCCGAGAGCTTAC 180
Db      219 GAGCTGAGACTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCCGAGAGCTTAC 277
QY      181 CGGCGCCGACGCGGCTTAAAGACTCTGCTGAGTGAAGCGCGCGGCACTGCGCGAG 240
Db      278 CGGCGCCGACGCGGCTTAAAGACTCTGCTGAGTGAAGCGCGCGGCACTGCGCGAG 337
QY      241 AGCAACTGCGGCTGCTGCTGCGCAACTCTGCGCGCTGCTGCGCGCACTGCTGCG 300
Db      338 AGCAACTGCGGCTGCTGCTGCGCAACTCTGCGCGCTGCTGCGCGCACTGCTGCG 397
QY      301 CACCTGCGCGCAAGCGGCGCGGCACTGCTCTCAGAAAGCTATAGCTATGCACTCC 360
Db      398 CACCTGCGCGCAAGCGGCGCGGCACTGCTCTCAGAAAGCTATAGCTATGCACTCC 455
QY      361 AGCTCTTCAAAGAGAGAGAGAGTACTGCTGCGCGCTGCGCACTGCTGCA 420
Db      456 AGCTCTTCAAAGAGAGAGAGAGTACTGCTGCGCGCTGCGCACTGCTGCA 515
QY      421 AATTCTCAGCAGGCTCAGTGGAGACAGGCTTCCCGCAACCAAGCGGCGGAGT 480
Db      516 AATTCTCAGCAGGCTCAGTGGAGACAGGCTTCCCGCAACCAAGCGGCGGAGT 575
QY      481 CGGCGCGCGGCTGCTGCTGCGCAAGCGGCGGAGAGGCGCGCGCAAGCGGCG 540
Db      576 CGGCGCGCGGCTGCTGCTGCGCAAGCGGCGGAGAGGCGCGCGCAAGCGGCG 635
QY      541 CAGCAGTGAAGCCCGCCAGACTTCTCTGAAAGCAAGTGAAGCTGTGATCCGCTC 600
Db      636 CAGCAGTGAAGCCCGCCAGACTTCTCTGAAAGCAAGTGAAGCTGTGATCCGCTC 695
QY      601 CGGCTTGAAGAGTACTGCGAGATGAGGCGGCTTGAAGCAGGCGGCTGCTCCGG 660
Db      696 CGGCTTGAAGAGTACTGCGAGATGAGGCGGCTTGAAGCAGGCGGCTGCTCCGG 755
QY      661 CGGCGCGGCGGCTGCTGCTGCGGAGCTGAGCGGCTTGGGCGGCGCAAGCTGCGC 720
Db      756 CGGCGCGGCGGCTGCTGCTGCGGAGCTGAGCGGCTTGGGCGGCGCAAGCTGCGC 815
QY      721 TCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      816 TCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
QY      781 GCTTCTT- GGGGCGACTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 839
Db      876 GCTTCTT- GGGGCGACTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 935
QY      840 GACTGAGGCGCTGAGAGAGCTGTGAGGCGGAGGCTGTGAGGCTGTGAGGCTGT 899
Db      936 GACTGAGGCGCTGAGAGAGCTGTGAGGCGGAGGCTGTGAGGCTGTGAGGCTGT 994
QY      900 TGAGG 904
Db      995 TGAGS 999

```

```

RESULT 6
LOCUS      BX384448
DEFINITION BX384448 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
ACCESSION  BX384448
VERSION     BX384448.2
KEYWORDS    EST.

```

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Li, W.B., Gruber, C., Jassse, J., and Polayes, D.
AUTHORS      Full-length cDNA libraries and normalization
TITLE        Unpublished (2001)
JOURNAL      On May 8, 2003 this sequence version replaced gi:30440338.
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen. This sequence belongs to sequence cluster
              9074.r
              For more information about this cluster, see
              http://www.genoscope.cns.fr/cdna?c=CS0DK010BF080P1&c=9074.r.
FEATURES
  source
    1..1021
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0DK010YK16"
    /cell_line="HELA"
    /clone_id="HELA"
    /note="1st strand cDNA was primed with a NotI-oligo(dT)
    primer. Five prime end enriched, double-strand cDNA was
    digested with Not I and cloned into the Not I and EcoR V
    sites of the pCMVSPORT 6 vector. Library was normalized."
  ORIGIN
    Query Match      90.5%; Score 822.8; DB 5; Length 1021;
    Best Local Similarity 97.0%; Pred. No. 1.2e-170;
    Matches 877; Conservative 12; Mismatches 9; Indels 6; Gaps 5;
QY      1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGATGAGTGCCTGGACTAC 60
Db      71 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGATGAGTGCCTGGACTAC 130
QY      61 TACGGGATGCTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 120
Db      131 TACGGGATGCTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 190
QY      121 GAGCTGAGACTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCCGAGAGCTTAC 180
Db      191 GAGCTGAGACTCTGCTGCTTCTGCTGATGAGGCTCTGCGCGCCGAGAGCTTAC 249
QY      181 CGGCGCGGCGGCTGCTGCTGCGCAAGCGGCTTGGGCGGCGCAAGCTGCGCGAG 240
Db      250 CGGCGCGGCGGCTGCTGCTGCGCAAGCGGCTTGGGCGGCGCAAGCTGCGCGAG 309
QY      241 AGCAACTGCGGCTGCTGCTGCGCAACTCTGCGCGCTGCTGCGCGCGCAAGCTGT 300
Db      310 AGCAACTGCGGCTGCTGCTGCGCAACTCTGCGCGCTGCTGCGCGCGCAAGCTGT 369
QY      301 CACCTGCGCGCAAGCGGCGCGGCACTGCTCTCAGAAAGCTATAGCTATGCACTCC 360
Db      370 CACCTGCGCGCAAGCGGCGCGGCACTGCTCTCAGAAAGCTATAGCTATGCACTCC 427
QY      361 AGCTCTTCAAAGAGAGAGAGTACTGCTGCGCGCTGCGCACTGCTGCAAGCACTTGC 419
Db      428 AGCTCTTCAAAGAGAGAGAGTACTGCTGCGCGCTGCGCACTGCTGCAAGCACTTGC 487
QY      420 AATTCTCAGCAGGCTCAGTGGAGACAGGCTTCCCGCAACCAAGCGGCGGAG 479
Db      488 AATTCTCAGCAGGCTCAGTGGAGACAGGCTTCCCGCAACCAAGCGGCGGAG 547
QY      480 TCGGCGCGGCGGCTGCTGCTGCGCAAGCGGCGGAGAGAGGCGCGCAAGCGGCG 539

```

Db 548 TCGGGGCGGCGCCAGAGTGTGTGTCAGACGCGCGGAGAGGGCGCCACCGACCGACCGCA 607
QY 540 GCAGCAGTCAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGCCTGTGACATCCGGCT 539
Db 608 GCAGCAGTCAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGCCTGTGACATCCGGCT 667
QY 600 CCGGGTTTCAGACAGTACTGTCGAGCATGGGCGCAGCTTGGAGCAGGGCGCTGGCATCCG 659
Db 668 CCGGGTTTCAGACAGTACTGTCGAGCATGGGCGCAGCTTGGAGCAGGGCGCTGGCATCCG 727
QY 660 GCGGCGCCAGGCGCTGGGCGGCGCAGTGAAGTGTGGGCGAGCGCCAGCATGTCGCG 719
Db 728 GCGGCGCCAGGCGCTGGGCGGCGCAGTGAAGTGTGGGCGAGCGCCAGCATGTCGCG 787
QY 720 CTCAGAGGACCTGGGCTCTGTGTTTGTGACATCAAGTCTTCAGAGCTCTCATCTGA 779
Db 788 CTCAGAGGACCTGGGCTCTGTGTTTGTGACATCAAGTCTTCAGAGCTCTCATCTGA 847
QY 780 CGCCTTCTGGGCGCAGTACTGATGAGCGCCCTGCTGACAGCGCTCGGGGCGTGTCT 839
Db 848 CGCCTTCTGGGCGGAGTGTGAGTGGCGCGCTGCTCAAGCCCTCGGGCGGCTGTCT 906
QY 840 GACTGAGCCCTCTGAGAGGCTGTGGGCGGAGGCTGTGCTGCTGTCAGTGTGA 899
Db 907 GACTGAGCCCTCTGAGAGGCTGTGGGCGGAGGCTGTGCTGCTGTCAGTGTGA 965
QY 900 TGAG 903
Db 966 TRAG 969

RESULT 7
BX374729
LOCUS BX374729 1091 bp mRNA linear EST 23-APR-2004
DEFINITION BX374729 Homo sapiens NEUROBLASTOMA COR 25-NORMALIZED Homo sapiens
CDNA clone CS0DC001YCI8 5-PRIME, mRNA sequence.
ACCESSION BX374729
VERSION BX374729.2 GI:46557336
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1091)
AUTHORS Li, W. B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30438519.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creteil, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DC001BB90P1&c=9074.r.

FEATURES

source

1. 1091

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DC001YCI8"

/issue_type="NEUROBLASTOMA COR 25-NORMALIZED"

/clone_id="Homo sapiens NEUROBLASTOMA COR 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 88.2%; Score 802; DB 5; Length 1091;
Best Local Similarity 96.8%; Pred. No. 4,7e-166;
Matches 881; Conservative 10; Mismatches 11; Indels 8; Gaps 7;

1 ATGGCCCTATCCGGGTGACACCCCGGCGCTGTGTGGAGAGAGATAGTCCCTGACTAC 60
Db 89 ATGGCCCTATCCGGGTGACACCCCGGCGCTGTGTGGAGAGAGATAGTCCCTGACTAC 148
QY 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGGCGGGCAATGACCGAGTGC 120
Db 149 TACGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGGCGGGCAATGACCGAGTGC 208
QY 121 GAGCTGAGCTCTGTGCTTCTTGTGATGAGAGCTCTTGGCGCGCCGAGGCTTAAGCC 180
Db 209 GAGCTGAGCTCTGTGCTTCTTGTGATGAGAGCTCTTGGCGCGCCGAGGCTTAAGCC 267
QY 181 CCGGCGCGCAGCGGCTTAAGAGCTCTGTGAGAGTGTGGAGCGCGCGGAGAGTGGCGAG 240
Db 268 CCGGCGCGCAGCGGCTTAAGAGCTCTGTGAGAGTGTGGAGCGCGCGGAGAGTGGCGAG 327
QY 241 AGCAACCTGCGGCTGTGGGCAACTCTGCGGCTGTGGCCCGGCAAGCTGTGCGC 300
Db 328 AGCAACCTGCGGCTGTGGGCAACTCTGCGGCTGTGGCCCGGCAAGCTGTGCGC 387
QY 301 CACTTGGCGGCAAGCGGCGCGGCGCAAGTGTCTCAAGAACCTTAAGTGAAGCACTTCC 360
Db 388 CACTTGGCGGCAAGCGGCGCGGCGCAAGTGTCTCAAGAACCTTAAGTGAAGCACTTCC 445
QY 361 AGCTTTCAAGAGAGACAGAGGTTAGCTGCGGTGCGGCGGAGTCAAGCAATTTCTGCA 420
Db 446 AGCTTTCAAGAGAGACAGAGGTTAGCTGCGGTGCGGCGGAGTCAAGCAATTTCTGCA 505
QY 421 AATTCTCAGCAGGCTCAGTGGAGAGCAGGCTCCCCCAACCAAGCGGCGGCGGAGT 480
Db 506 AATTCTCAGCAGGCTCAGTGGAGAGCAGGCTCCCCCAACCAAGCGGCGGCGGAGT 565
QY 481 CCGGCGCGGCGCAGTGGTGTGCAGACGCGGCGGAGAGGCGCCCAAGCTGAGCTCCAG 540
Db 566 CCGGCGCGGCGCAGTGGTGTGCAGACGCGGCGGAGAGGCGCCCAAGCTGAGCTCCAG 624
QY 541 CAGCAGTCAGAGCGCCCGCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCGGCTC 600
Db 625 CAGCAGTCAGAGCGCCCGCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCGGCTC 684
QY 601 CCGGTTTCAGACAGTACTGTCGAGCATGGGCGCAGCTTGGAGCAGGAGCGTGCATCCCG 660
Db 685 CCGGTTTCAGACAGTACTGTCGAGCATGGGCGCAGCTTGGAGCAGGAGCGTGCATCCCG 744
QY 661 CCGCGCCAGGCGCTGGCGCGGCGAGCTGAGCGTGTGGGCGAGGCAACCGAGTGTGCGC 720
Db 745 CCGCGCCAGGCGCTGGCGCGGCGAGCTGAGCGTGTGGGCGAGGCAACCGAGTGTGCGC 802
QY 721 TCAAGGACCTGGGCTGTGTGTGTGTGACATGAATTTCTCAGAGCTCTCTCATCTGAGC 780
Db 803 TCAAGGACCTGGGCTGTGTGTGTGTGACATGAATTTCTCAGAGCTCTCTCATCTGAGC 862
QY 781 GCTTCTGAGGAGTACTGAGTGGCGGCTGTGACAGGCGCTGCGGGGCGTGTCTCT 839
Db 863 GCTTCTGAGGAGTACTGAGTGGCGGCGGCTGTGACAGGCGCTGCGGGGCGTGTCTCT 922
QY 840 GACTGAGCCCTCTGAGAGGCTGTGGGCGGAGGAGCTGTTGCGCTGTGTCAGTGTGA 899
Db 923 GACTGAGCCCTCTGAGAGGCTGTGGGCGGAGGAGCTGTTGCGCTGTGTCAGTGTGA 981
QY 900 TGAGGCTGAC 909
Db 982 TGAGGCTGAC 991

RESULT 8
BX375652

LOCUS BX375652 951 bp mRNA linear EST 26-APR-2004
 DEFINITION BX375652 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 CDNA clone CS0DC015YJ24 5'-PRIME, mRNA sequence.
 ACCESSION BX375652
 VERSION BX375652.2 GI:46573220
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 951)
 AUTHORS Li, W.B., Gruber, C., Jeesee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On May 8, 2003 this sequence version replaced gi:30448429.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9074.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdnas?CS0DC015DE12QPI&c=9074.r.
 FEATURES
 SOURCE
 1. 951
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC015YJ24"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 85.1%; Score 773.6; DB 5; Length 951;
 Best Local Similarity 98.0%; Pred. No. 8.3e-160;
 Matches 833; Conservative 2; Mismatches 8; Indels 7; Gaps 5;

1 ATGCGGATCGGGTGCACCCCGCGCTGGGAGAGATGAGTGCCTGACAC 60
 76 ATGCGGATCGGGTGCACCCCGCGCTGGGAGAGATGAGTGCCTGACAC 135
 61 TACGGAGTGTGCTGCCTTCACTGATGTTGAGGTGTGGCGGCACTGACGAGTGC 120
 136 TAGGGAGTGTGCTGCCTTCACTGATGTTGAGGTGTGGCGGCACTGACGAGTGC 195
 121 GAGCTGAGTGTGCTGCCTTCTGCTGATGAGTGCCTGCGCGCGGAGGCTTAC 180
 196 GAGCTGAGTGTGCTGCCTTCTGCTGATGAGTGCCTGCGCGCGGAGGCTTAC 254
 181 CGGGCCGCGAGCGCTTGAAGTCTGCTGAGTGTGGAGCGCGCGGAGTGTGGAG 240
 255 CGGGCCGCGAGCGCTTGAAGTCTGCTGAGTGTGGAGCGCGCGGAGTGTGGAG 314
 241 AGCAACCTGCGGCTGTGGGCAACTCTGCGCGTGTGGCGCGGCACTGCGCGG 300
 315 AGCAACCTGCGGCTGTGGGCAACTCTGCGCGTGTGGCGCGGCACTGCGCGG 374
 301 CACTGTGGCGCGCAAGCGCGCGGCACTGTCTTCAGAACTATAGTATGACCTCC 360
 375 CACTGTGGCGCGCAAGCGCGCGGCACTGTCTTCAGAACTATAGTATGACCTCC 432
 361 AGCTCTTCAAGAGAGACAGAGGTAGTGCCTGCGCGCTGCGGCACTGAGAGTTTCA 420
 433 AGCTCTTCAAGAGAGACAGAGGTAGTGCCTGCGCGCTGCGGCACTGAGAGTTTCA 492

421 AATTCTAGCAGGGTCACTGAGAGACAGGCTCCCCCAACCAAGCGGACCG- GCCGAG 479
 493 AATTCTAGCAGGGTCACTGAGAGACAGGCTCCCCCAACCAAGCGGACCGGAG 552
 480 TCGGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 539
 553 TCGGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 612
 540 GCACAGTCAAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 599
 613 GCACAGTCAAGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 672
 600 CCGGGTTCAGACAGAGTACTGCGAGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659
 673 CCGGGTTCAGACAGAGTACTGCGAGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 732
 660 GCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719
 733 GCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 792
 720 CTCAGAGGACCTGAGGCTCTGCTGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGA 779
 793 CTCAGAGGACCTGAGGCTCTGCTGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGA 851
 780 GCGCTTCTGAGGCGAGTCACTGAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGG 839
 852 GCGCTTCTGAGGCGAGTCACTGAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 909
 840 GACTGAGGCC 849
 910 GACTGAGGCC 919

RESULT 9
 LOCUS BG685173 800 bp mRNA linear EST 01-MAY-2001
 DEFINITION BG685173 NIH_MGC_48 Homo sapiens CDNA clone IMAGE:4764639 5',
 mRNA sequence.
 ACCESSION BG685173
 VERSION BG685173.1 GI:13916570
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 800)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strauberg, Ph.D.
 Email: sgabs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM1622 row: d column: 16
 High quality sequence stop: 794.
 Location/Qualifiers
 1. 800
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4764639"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 48"
 /note="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCAAGAG(G). Size-selected >500bp

ORIGIN

for average insert size 1.8kb. Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 83.7%; Score 760.4; DB 4; Length 800;
Best Local Similarity 99.6%; Pred. No. 6.6e-157;
Matches 783; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

119 GCGAGCTGAGCTCTGCGCTCTTCTGCTGATGAGCTCTGCGCGCGCGAGGCTTAG 178
|||
2 GCGAGCTGAGCTCTGCGCTCTTCTGCTGATGAGCTCTGCGCGCGCGAGGCTTAG 61
|||
179 CCGCGCGCGCGAGCTGAGCTCTGCTGAGCTGAGCGCGCGCGAGTGCAGC 238
|||
62 CCGCGCGCGCGAGCTGAGCTCTGCTGAGCTGAGCGCGCGCGAGTGCAGC 120
|||
239 AGAGCAACTGCGGCTGCTGCGGCACTCTGCGCGCTGCGCGCGAGCTGCTGC 238
|||
121 AGAGCAACTGCGGCTGCTGCGGCACTCTGCGCGCTGCGCGCGAGCTGCTGC 180
|||
299 CGGACCTGCGGCGCGAGCGCGCGCGAGCTCTGCGCGCTGAGTATGAGCACT 358
|||
181 CGGACCTGCGGCGCGAGCGCGCGCGAGCTCTGCGCGCTGAGTATGAGCACT 240
|||
359 CGGACCTGCGGCGCGAGCGCGCGCGAGCTCTGCGCGCTGAGTATGAGCACT 418
|||
241 CGGACCTGCGGCGCGAGCGCGCGCGAGCTCTGCGCGCTGAGTATGAGCACT 300
|||
419 CAATTTCTGAGAGGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 478
|||
301 CAATTTCTGAGAGGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 360
|||
479 GTGCGGCGCGCGCGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 538
|||
361 GTGCGGCGCGCGCGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 420
|||
539 AGGAGCTGAGAGCGCGCGAGCTCTGCGCGCGAGCGCGAGTGCAGTGCAG 598
|||
421 AGGAGCTGAGAGCGCGCGAGCTCTGCGCGCGAGCGCGAGTGCAGTGCAG 480
|||
599 TCCGCGGCTGAGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 658
|||
481 TCCGCGGCTGAGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 540
|||
659 GCGCGCGCGCGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 718
|||
541 GCGCGCGCGCGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 600
|||
719 GCTCAAGGAGCTGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 778
|||
601 GCTCAAGGAGCTGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 660
|||
779 AGCGCTTCTGAGGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 838
|||
661 ACCGCTTCTGAGGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 719
|||
839 TGACTGAGGCTGAGGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 898
|||
720 TGACTGAGGCTGAGGAGCTGAGTGGAGAGAGCTCTGCGCGCGAGCGCGAGTGCAG 779
|||
899 ATGAGG 904
|||
780 ATGAGG 785

RESULT 10
B1769587
LOCUS B1769587 837 bp mRNA linear EST 25-SEP-2001
DEFINITION 60305496.F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204607 5',
mRNA sequence.

ACCESSION B1769587
VERSION B1769587.1 GI:15761165
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 837)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM1512 row: P column: 16
High quality sequence stop: 830.

FEATURES
source
1..837
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5204607"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source: anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 026. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 81.2%; Score 738.4; DB 4; Length 837;
Best Local Similarity 99.2%; Pred. No. 4.6e-152;
Matches 794; Conservative 0; Mismatches 1; Indels 5; Gaps 5;

1 ATGGCGCTATCCGGGTCGAGACCCCGCGCGCTGAGGAGAGATGCTGACTAC 60
|||
40 ATGGCGCTATCCGGGTCGAGACCCCGCGCGCTGAGGAGAGATGCTGACTAC 99
|||
61 TACGGGATGCTGCTTCAACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 120
|||
100 TACGGGATGCTGCTTCAACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 159
|||
121 GAGCTGAGTCTCTGCGCTTTTCTGCTGAGTGAAGCTCTGCGCGCGCGAGCTTAA 180
|||
160 GAGCTGAGTCTCTGCGCTTTTCTGCTGAGTGAAGCTCTGCGCGCGCGAGCTTAA 219
|||
181 CCGGCGCGCGAGCGGCTGAGAGCTCTGCTGAGAGTGAAGCGCGCGAGTGCAG 240
|||
220 CCGGCGCGCGAGCGGCTGAGAGCTCTGCTGAGAGTGAAGCGCGCGAGTGCAG 279
|||
241 AGCAACTGCGGCTGCTGAGGCAACTCTGCGGCTGAGCGCGCGCACTGCTGCG 300
|||
280 AGCAACTGCGGCTGCTGAGGCAACTCTGCGGCTGAGCGCGCGCACTGCTGCG 339
|||
301 CACTGCGCGCGAGCGCGCGCGCGCGAGTGTCTCGAAGAGCTATGACTATGCACTCC 360
|||
340 CACTGCGCGCGAGCGCGCGCGCGCGAGTGTCTCGAAGAGCTATGACTATGCACTCC 399
|||
361 AGCTTTCAAGAGGACAGAGGAGTAGCTGCGCGCGCTGCGCGAGTCAAGCACTTCA 420
|||
400 AGCTTTCAAGAGGACAGAGGAGTAGCTGCGCGCGCTGCGCGAGTCAAGCACTTCA 459
|||

QY 421 AATTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
 DB 460 AATTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 518
 QY 481 CGGGGCGGGCCAGTGTGTGTGTCCAGACGCGGCGGAGAGGGGCGCCAGCCCA-CCCCA 539
 DB 519 CGGGGCGGGCCAGTGTGTGTGTCCAGACGCGGCGGAGAGGGGCGCCAGCCCA-CCCCA 577
 QY 540 GCACAGTCAGAGCCCGGACAGACCTTCTCTGAAGGCAAGAGTGAATCTGATCCGAGCT 599
 DB 578 GCACAGTCAGAGCCCGGACAGACCTTCTCTGAAGGCAAGAGTGAATCTGATCCGAGCT 637
 QY 600 CGGGGTTGAGCAGAGTACTGAGCA-TGGGCGGAGCTTGAAGAGGAGCGCT-GGCAATCC 657
 DB 638 CGGGGTTGAGCAGAGTACTGAGCA-TGGGCGGAGCTTGAAGAGGAGCGCTGAGCAATCC 697
 QY 658 CGGGGCGGGCCAGTGTGTGTGTCCAGACGCGGCGGAGAGCTGTGTGGGAGCGCAAGCGGAGTGTG 717
 DB 698 CGGGGCGGGCCAGTGTGTGTGTCCAGACGCGGCGGAGAGCTGTGTGGGAGCGCAAGCGGAGTGTG 757
 QY 718 CGCTCAAGGAGCAGTGTGTGTGTGTGTGACATCAAGTCTGAGAGCTCTCCATCTG 777
 DB 758 CGCTCAAGGAGCAGTGTGTGTGTGTGTGACATCAAGTCTGAGAGCTCTCCATCTG 817
 QY 778 GACGCTTCTGGGGGAGCTA 797
 DB 818 GACGCTTCTGGGGGAGCTA 837

RESULT 11
 BM925969 1089 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT_6649774 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764272
 DEFINITION 5', mRNA sequence.
 ACCESSION BM925969
 VERSION BM925969.1 GI:19376336
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1089)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LAM12817 row: h column: 01
 High quality sequence start: 14
 High quality sequence stop: 586.
 Location/Qualifiers

FEATURES

source

1. 1089
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5764272"
 /lab_host="DH10B"
 /clone_11b="NIH_MGC_114"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber

(Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

Query Match 80.0%; Score 727.4; DB 5; Length 1089;
 Best Local Similarity 95.7%; Pred. No. 1.2e-149;
 Matches 780; Conservative 0; Mismatches 31; Indels 4; Gaps 3;

ORIGIN
 1 ATGCGGATATCCGGGTGACCCCGGCGGCTGCTGGAGAGAGATGATGCTGACTAC 60
 DB 71 ATGCGGATATCCGGGTGACCCCGGCGGCTGCTGGAGAGAGATGATGCTGACTAC 130
 QY 61 TAGGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGTGGGCAATCTGACGAGTGC 120
 DB 131 TAGGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGTGGGCAATCTGACGAGTGC 190
 QY 121 GAGCTGAGTCTTGTGCTTCTGCTGATGAGCTTCTGAGCGGCGGAGGCTTATGCC 180
 DB 191 GAGCTGAGTCTTGTGCTTCTGCTGATGAGCTTCTGAGCGGCGGAGGCTTATGCC 250
 QY 181 CGGGCGGAGCGGCTTGAAGCTCTGTGAGTGTGAGCGCGGAGGAGTGTGCGAG 240
 DB 251 CGGGCGGAGCGGCTTGAAGCTCTGTGAGTGTGAGCGCGGAGGAGTGTGCGAG 310
 QY 241 AGCAACCTGTGCTGTGGGCAATCTCTGTGCTGTGCGCGGCGGCAAGACTGTGCGG 300
 DB 311 AGCAACCTGTGCTGTGGGCAATCTCTGTGCTGTGCGCGGCGGCAAGACTGTGCGG 370
 QY 301 CACTGTGCGGCAAGCGGCGGCGGCGGCTTCTCAAGAGCTTATGCTATGCACTGCC 360
 DB 371 CACTGTGCGGCAAGCGGCGGCGGCGGCTTCTCAAGAGCTTATGCTATGCACTGCC 430
 QY 361 AGCTTTCAAGAGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCAAGAGCTTGTGCA 420
 DB 431 AGCTTTCAAGAGAGAGAGAGAGTGTGCGGCGGCGGCGGCAAGAGCTTGTGCA 490
 QY 421 AATTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
 DB 491 AATTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 550
 QY 481 CGGGGCGGGCCAGTGTGTGTGTCCAGACGCGGCGGAGAGGGGCGCCAGCCCGAG 540
 DB 551 CGGGGCGGGCCAGTGTGTGTGTCCAGACGCGGCGGAGAGGGGCGCCAGCCCGAG 610
 QY 541 CAGCAGTCAGAGCCCGGAGACCTTCTCTGAAGGCAAGTGAATCTGATCCGAGTGC 600
 DB 611 CAGCAGTCAGAGCCCGGAGACCTTCTCTGAAGGCAAGTGAATCTGATCCGAGTGC 670
 QY 601 CGGGTTCAGAGAGTACTGAGCATGAGGCGGAGCTTGTGAGCAGGCGTGTGATCCCGG 660
 DB 671 CGGGTTCAGAGAGTACTGAGCATGAGGCGGAGCTTGTGAGCAGGCGTGTGATCCCGG 730
 QY 661 CGGGCGGAGCGGCTGTGCGGCGGAGCTGTGAGC-TGTTTGGCAGGCGCAAGCGAGTGTG-CTGC 718
 DB 731 GGGGCGGAGCGGCTGTGCGGCGGAGCTGTGAGC-TGTTTGGCAGGCGCAAGCGAGTGTG-CTGC 790
 QY 719 GCTCAAGAGAGCTGTGCT--CTGTGTGTGTGATCATGAATCTCAAGCTCTTATCT 776
 DB 791 GCTCAAGAGAGCTGTGCT--CTGTGTGTGTGATCATGAATCTTCAAGCTCTTCAATT 850
 QY 777 GAGCGCTTCTGGGGGAGTACTGATGAGCGGCC 811
 DB 851 GAGCGCTTCTGGGGGAGTACTGATGAGCGGCC 885

RESULT 12
 BP158001 820 bp mRNA linear EST 30-DEC-2003
 LOCUS BP158001 full-length enriched swine cDNA library, adult spleen sus
 DEFINITION BP158001
 ACCESSION BP158001
 VERSION BP158001.1 GI:40407474
 KEYWORDS EST.

SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 820)
 AUTHORS Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,
 Okumura, N., Hamashima, N. and Awata, T.
 TITLE PEDF (Pig EST Data Explorer): construction of a database for ESTs
 derived from porcine full-length cDNA libraries
 Nucleic Acids Res. 32 (1), D484-D488 (2004)
 JOURNAL Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in
 Animal Genome Research Program (Japan) by National Institute of
 Agrobiological Sciences and STARP-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA
 library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13
 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.
 FEATURES
 source
 1..820
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="SPL010019D12"
 /tissue_type="spleen"
 /dev_stage="adult"
 /clone_lib="full-length enriched swine cDNA library, adult
 spleen"
 ORIGIN
 Query Match 78.2%; Score 711.2; DB 5; Length 820;
 Best Local Similarity 92.6%; Pred. No. 4.5e-146;
 Matches 759; Conservative 0; Mismatches 58; Indels 3; Gaps 1;
 44 ATGAGTCTGTGACTACTACGGGATGCTGTGCTTCAACCGATGTTGAGAGTGGGCG 103
 1 ATGAGTCTGTGACTACTACGGGATGCTGTGCTTCAACCGATGTTGAGAGTGGGCG 60
 104 GGCAACTGACGAGTGGAGCTGAGCTCTGGCTTTCTGTGATGAGAGCTCTGGCG 163
 61 GGCAACTGAGTGGAGCTGAGCTCTGGCTTTCTGTGATGAGAGCTCTGGCG 120
 164 CCGCCGAGGCTTACCCCGGCGCGAGCGGCTTACGCTCTGCTGAGTGGAGCGC 223
 121 CCGCCGAGGCTTACCCCGGCGCGAGCGGCTTACGCTCTGCTGAGTGGAGCGC 180
 224 GGCGGCGAGTGGGAGAGCACTGCGGCTGCGGGCAACTCTGCGGCTGCGGCC 283
 181 GGCGGCGAGTGGGAGAGCACTGCGGCTGCGGGCAACTCTGCGGCTGCGGCC 240
 284 GGCAAGCACTGTGCGCACCTGCGCGGCAAGCGCGCCGCGCAAGTCTTCAAGAGCT 343
 241 GGCAAGCACTGTGCGCGCACCTGCGCGGCAAGCGCGCCGCGCAAGTCTTCAAGAGCT 300
 344 ATGAGTATGCACTTCAAGCTCT--TCAAGAGAGAGAGGTAAGTCTGCGGCTGCGGTC 400
 301 ACAGCTATGCACTTCAAGCTCTTCTTCAAGAGAGAGTCTGCGGCTGCGGTC 360
 401 GGAGTCAACAGCTTCAAGCTCTTCTTCAAGAGAGAGTCTGCGGCTGCGGTC 460
 361 GGAGTCAACAGCTTCTTCAAGCTCTTCAAGAGAGAGTCTGCGGCTGCGGTC 420
 461 CCAAGCGGAGCGCGAGTCTGCGGCGCGGCGAGTCTGCGGCGAGCGCGGCGAGAG 520
 421 CCAAGCGGAGCGCGAGTCTGCGGCGCGGCGAGTCTGCGGCGAGCGCGGCGAGAG 480

QY 521 GGCGCCAGCGGAGCGGAGCACTGAGAGCGCGGAGAGCTTCTCTGAAGGCAAG 580
 DB 481 GGCGCCAGCGGAGCGGAGCGGAGCACTGAGAGCGCGGAGAGCTTCTCTGAAGGCAAG 540
 QY 581 TGACCTGTGACATCCCGGCTCCGGGTTTGAGCAGAGTACTGAGCATGAGCGGCGCTTGG 640
 DB 541 TGACCTGTGATATCCGGCTCAGGGGTGAGCAGAGTACTGATGATGAGGCGGCGCTTGG 600
 QY 641 AGAGAGGCGTGTGATCCCGGCGGCGGAGCGGCGGCGGAGCTGAGAGTCTTGGGC 700
 DB 601 AGCAGGCGGTGTGATCCCGGCGGCGGAGCGGCGGAGCTGAGAGTCTTGGGC 660
 QY 701 AGGCGACCGAGTGTGCGGCTCAAGGAGTGGGCTGTGTTGTGATCAAGTCTT 760
 DB 661 AGGCTACAGAGTGTGCGGCTCAAGGAGTGGGCTGTGTTGTGATCAAGTCTT 720
 QY 761 CAGAGCTTCTTATCTGACAGCGCTTCTGAGGCGGAGTACTGAGTGTGCGGCTGCTCAGG 820
 DB 721 CAGAGCTTCTTATCTGAGAGCGCTTCTGAGGCGGAGTACTGAGTGTGCGGCTGCTCAGG 780
 QY 821 CCTTGGCGGCGGTGTCTTCTGACTGAGGCGGCTGCGAGAGC 860
 DB 781 CCTTGGCGGCGGTGTCTTCTGACTGAGGCGGCTGCGAGAGC 820
 RESULT 13
 BM920838
 LOCUS 1090 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT 6706034 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5752092
 5', mRNA Sequence.
 ACCSSION BM920838
 VERSION BM920838.1 GI:19371217
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1090)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL2785 row: 1 column: 13
 High quality sequence stop: 625.
 FEATURES
 source
 1..1090
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5752092"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 77.5%; Score 704.2; DB 5; Length 1090;

Best Local Similarity 95.5%; Pred. No. 1.6e-144;
Matches 780; Conservative 0; Mismatches 28; Indels 9; Gaps 5;

```
QY 1 ATGGCGGTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTCTGCTGACTAC 60
DB 102 ATGGCGGTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTCTGCTGACTAC 161
QY 61 TACGGAGTGTGTGCTTCAACCGATGTTGAGGTGAGGCGGAGCAACTGACGAGTGC 120
DB 162 TACGGAGTGTGTGCTTCAACCGATGTTGAGGTGAGGCGGAGCAACTGACGAGTGC 221
QY 121 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAC 180
DB 222 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAC 281
QY 181 CGGGCGCGGAGCGGCTTGAAGCTCTGCTGAGAGCGCGCGCGGCACTGAGGCGAG 240
DB 282 CGGGCGCGGAGCGGCTTGAAGCTCTGCTGAGAGCGCGCGCGGCACTGAGGCGAG 341
QY 241 AGCAACTGCGGCTGCTGAGGCAACTCTGCTGCGGCTGAGCGCGCAAGCTGCTGCG 300
DB 342 AGCAACTGCGGCTGCTGAGGCAACTCTGCTGCGGCTGAGCGCGCAAGCTGCTGCG 401
QY 301 CACCTGCGCGCAAGCGCGCGCGCAAGTGTCTCAGAAAGCTATAGCTATGCACTCC 360
DB 402 CACCTGCGCGCAAGCGCGCGCGCAAGTGTCTCAGAAAGCTATAGCTATGCACTCC 461
QY 361 AGCTCTTCAAGAGAGACAGAGGTTAGTGCCTGCTGCGGCTGAGAGAGTTTCA 420
DB 462 AGCTCTTCAAGAGAGACAGAGGTTAGTGCCTGCTGCGGCTGAGAGAGTTTCA 521
QY 421 AATTCTAGACAGGTTCAAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
DB 522 AATTCTAGACAGGTTCAAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 581
QY 481 CGGGCGCGGCGCAAGTGTGTGCTGAGAGCGGCGGAGAGGCGCGCAAGCGGCA 540
DB 582 CGGGCGCGGCGCAAGTGTGTGCTGAGAGCGGCGGAGAGGCGCGCAAGCGGCA 641
QY 541 CAGAGTCAAGAGCGCGCGCAAGCTTCTCTTCAAGCAAGTGAATCTGATCCGCTC 600
DB 642 CAGAGTCAAGAGCGCGCGCAAGCTTCTCTTCAAGCAAGTGAATCTGATCCGCTC 701
QY 601 CGGGTTGAGAGAGTACTGAGAGATGAGGCGGCTTGAAGAGAGGCGGATCCCG 660
DB 702 CGGGTTGAGAGAGTACTGAGAGATGAGGCGGCTTGAAGAGAGGCGGATCCCG 761
QY 661 CGGGCGGAGGCTGAGGCGGCGGAGCTGAGCGTGT - GGGCAGGCGCAAGCGAGTGTG- 717
DB 762 CGGGCGGAGGCTGAGGCGGCGGAGCTGAGCGTGT - GGGCAGGCGCAAGCGAGTGTG- 821
QY 718 CGCTCAAGAGAGCTGAGGCTGAGGCTTGTGAGTGTGAGT - CAAGTGTCTAGAG- -CTCTCC 771
DB 822 CGCTCAAGAGAGCTGAGGCTTGTGAGTGTGAGTGTGAGTGTCTCAAGAGCTCTCCCT 881
QY 772 TATCTGAGAGGCTCTGAGGCGGAGTCACTGAGTGGCG 808
DB 882 ATCTGAGAGGCTCTGAGGCGGAGTCACTGAGTGGAG 918
```

RESULT 14
BI523561
LOCUS 603175851.F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240077 5',
DEFINITION mRNA sequence.
ACCESSION BI523561
VERSION BI523561.1 GI:15348353
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 872)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straube, Ph.D.
Email: egads-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM11605 row: F column: 14
High quality sequence start: 28
High quality sequence stop: 805.
Location/Qualifiers

FEATURES
source
1..872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5240077"
/lab_host="DH10B"
/clone_id="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 74.9%; Score 680.8; DB 4; Length 872;
Best Local Similarity 98.8%; Pred. No. 2.2e-139;
Matches 759; Conservative 0; Mismatches 2; Indels 7; Gaps 7;

```
QY 1 ATGGCGGTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTCTGCTGACTAC 60
DB 108 ATGGCGGTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTCTGCTGACTAC 167
QY 61 TACGGAGTGTGTGCTTCAACCGATGTTGAGGTGAGGCGGAGCAACTGACGAGTGC 120
DB 168 TACGGAGTGTGTGCTTCAACCGATGTTGAGGTGAGGCGGAGCAACTGACGAGTGC 227
QY 121 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAC 180
DB 228 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAC 287
QY 228 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAC 347
DB 288 CGGGCGGAGGCTTGAAGCTCTGCTGAGAGCGGCGGAGGCGGAGGCTTGAAGGAG 407
QY 288 CGGGCGGAGGCTTGAAGCTCTGCTGAGAGCGGCGGAGGCGGAGGCTTGAAGGAG 467
DB 348 AGCAACTGCGGCTGCTGAGGCAACTCTGCTGCGGCTGAGCGCGCAAGCTGCTGCG 527
QY 348 AGCAACTGCGGCTGCTGAGGCAACTCTGCTGCGGCTGAGCGCGCAAGCTGCTGCG 587
DB 408 GCACTGCGGCTGCTGAGGCAACTCTGCTGCGGCTGAGCGCGCAAGCTGCTGCG 547
QY 408 GCACTGCGGCTGCTGAGGCAACTCTGCTGCGGCTGAGCGCGCAAGCTGCTGCG 607
DB 468 TCCAGCTCTTCAAGAGAGACAGAGGTTAGTGCCTGCTGCGGCTGAGCGCGCAAG 667
QY 468 TCCAGCTCTTCAAGAGAGACAGAGGTTAGTGCCTGCTGCGGCTGAGCGCGCAAG 727
DB 528 GCAATTTCTAGAGAGGTTCAAGAGGTTAGTGCCTGCTGCGGCTGAGCGCGCAAG 787
QY 528 GCAATTTCTAGAGAGGTTCAAGAGGTTAGTGCCTGCTGCGGCTGAGCGCGCAAG 847
DB 587 GAGTGGGGCGGCGGAGGTTGCTGAGAGCGGCGGAGGAGGAGGCGGAGGCGGAGG 947
QY 587 GAGTGGGGCGGCGGAGGTTGCTGAGAGCGGCGGAGGAGGAGGCGGAGGCGGAGG 1007
```

Db 588 GAGTCGGGGCCGGCCAGAGTGTGTCACAGACGGCGGAGAGGGGCCCGCCAGCCGA-C 646

QY 537 CCAGCAGACGACTCAGAGACCCCGCAGACCTTCTCTGAAAGCAAAAGTGAAGCTGTGACATCCG 536

Db 647 CCAGCAGACGACTCAGAGACCCCGCAGACCTTCTCTGAAAGCAAAAGTGAAGCTGTGACATCCG 705

QY 597 GCTCCGGGTTCCAGCAGAGTACTGCGAGCATGGGCCAGCTTGGAGAGAGGCGGTGGCATC 656

Db 706 GCTCCGGGTTCCAGCAGAGTACTGCGAGCATGGGCCAGCTTGGAGAGAGGCGGTGGCATC 765

QY 657 CCGGCGGCCCCCAGACCGCTGGCGCGGCGAGCTGACGTTGTTGGCGAGGCCAGCGAGTGT 716

Db 766 CCGGCGGCCCCCAGACCGCTGGCGCGGCGAGCTGACGTTGTTGGCGAGGCCAGCGAGTGT 825

QY 717 GCGCTCAGAGGAGCTGGGCTCTGTGTTGTGATCATCAAGTCTCAGA 764

Db 826 GGGCTCAGAGGAGCTGGGCTCTGTG-TGTGACATCAAGTCTCAGA 872

RESULT 15

AL519410 874 bp mRNA linear EST 24-MAR-2004

LOCUS AL519410 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens

DEFINITION cDNA clone CS0DB004YF09 5-PRIME, mRNA sequence.

ACCESSION AL519410

VERSION AL519410.3 GI:45694960

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 874)

AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.

TITLE Full-length cDNA library and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 13, 2001 this sequence version replaced gi:31037757.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

9074.i

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0DB004CC050P1&c=9074.i.

FEATURES

source

1..874

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DB004YF09"

/tissue="NEUROBLASTOMA COT 10-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 74.7%; Score 678.8; DB 1; Length 874;

Best Local Similarity 89.8%; Pred. No. 6,1e-139;

Matches 750; Conservative 36; Mismatches 42; Indels 7; Gaps 6;

QY 1 ATGGCGTATCCGGGTGAGACCCCGCGCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 60

Db 43 ATGGCGTATCCGGGTGAGACCCCGCGCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 102

QY 61 TAGCGATGCTGTCGCTTCAACCGATGTTGAGAGTGGCGGCAACTGACCGAGTGC 120

Db 103 TAGCGATGCTGTCGCTTCAACCGATGTTGAGAGTGGCGGCAACTGACCGAGTGC 162

QY 121 GAGCTGAGAGTCTCTGGCCTTTTCTGCTGATAGAGGCTCTTGGCGCCCGGAGGCTTAGCC 180

Db 163 GAGCTGAGAGTCTCTGGCCTTTTCTGCTGATAGAGGCTCTTGGCGCCCGGAGGCTTAGCC 221

QY 181 CCGGCGCGAGCGGCTAGAGCTCTGCTGAGCTGAGAGCGCGCGGCG -GAGTGGCGGCGA 239

Db 222 CCGGCGCGAGCGGCTTAGAGCTCTGCTGAGCTGAGAGCGCGCGGAGAGTGGCGAGCA 281

QY 240 GAGCAACCTGCGGCTCTGGGCGCAACTCTGCGGATGCTGGCGCGCAGCAAGCTGTGCC 299

Db 282 GAGCAACCTGCGGCTCTGGGCGCAACTCTGCGGATGCTGGCGCGCAGCAAGCTGTGCC 341

QY 300 GCACTGGCGCGGCAACCGCGCGCGGCGGAGTGTCTCAGAACGTTATAGCTATAGGCACTC 359

Db 342 GCACTGGCGCGGCAACG -GCGCGGCGGAGTGTCTCAGAACGTTATAGCTATAGGCACTC 400

QY 360 CAGCTCTTCAAGAGAGAGAGGAGTGTGCGGCTGCGCGGCGGAGTCAAGAGAGTGTGCG 419

Db 401 CAGCTCTTCAAGAGAGAGAGGAGTGTGCGGCTGCGCGGCGGAGTCAAGAGAGTGTGCG 460

QY 420 AAATTCTCAG -CAGGCTCAGTGGAGAGAGAGTCTCCCGCCCAACCAAGCGGAGCGGCGA 478

Db 461 AAATTCTCAGAGAGGAGTGTGAGAGAGAGTCTCCCGCCCAACCAAGCGGAGCGGCGA 520

QY 479 GTGCGGCGCGGCGGAGTGTGTGCTCAGAGAGCGGCGGAGAGAGGCGGCGGAGCGGCGG 538

Db 521 GTGCGGCGCGGCGGAGTGTGTGCTCAGAGAGCGGCGGAGAGAGGCGGCGGAGCGGCGG 580

QY 539 AGCAGAGTCAAGAGCGCGGCGGAGAGTCTTCTCTGAAAGCAAGTGAAGTGTGATCATCCG 598

Db 581 AGCAGAGTCAAGAGCGCGGCGGAGAGTCTTCTCTGAAAGCAAGTGAAGTGTGATCATCCG 640

QY 599 TCCGAGTTCAGAGAGTACTGAGACATGAGGCGAGCTTGGAGAGAGGCGGTGATCC 658

Db 641 TCCGAGTTCAGAGAGTACTGAGAGCGGCGGAGAGTCTTGGAGAGAGGCGGTGATCC 700

QY 659 GCGGCGCGGCGGAGTGTGCGGCGGCGGAGTGTGTTGGGCGAGGCGGAGTGTGCTGC 718

Db 701 GCGGCGCGGCGGAGTGTGCGGCGGCGGAGTGTGTTGGGCGAGGCGGAGTGTGCTGC 760

QY 719 GCTCAAGGAGAGTGTGCTGTGTTGTGACATCAAGTCTCTGAGAGTCTCTATCTGG 778

Db 761 GATCCCGGCGGAGTGTGTTGTGACATCAAGTCTCTGAGAGTCTCTATCTGG 820

QY 779 ACAGCTTCTGAGGAGAGTACTGAGTGGCGGCGGCTGTGAGAGGCGGCGGCGT 833

Db 821 MGG -GTTTGGGCGGCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 872

Search completed: March 23, 2005, 16:09:29

Job time : 3269.36 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 07:12:12 ; Search time 4091.8 Seconds
(without alignments)
10764.399 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctacccgggtcgcac.....tcagctgcatgagcgtcgcac 909

Scoring table: IDENTITY NUC
Gapop 10-0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_on: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	909	100.0	909	6	BD093311 Apoptosis
2	909	100.0	1883	6	BD093312 Apoptosis
3	909	100.0	1883	9	AK075328 Homo sapi
4	907.4	99.8	981	6	C0719642 Sequence
5	907.4	99.8	981	9	AF443591 Homo sapi
6	907.4	99.8	1230	6	AX322754 Sequence
7	907.4	99.8	1924	6	AX431308 Sequence
8	904.2	99.5	1951	9	BC027930 Homo sapi
9	899.4	98.9	981	9	AF457575 Homo sapi
10	865.8	95.2	2012	9	AY125488 Homo sapi
11	845.4	93.0	2005	9	BC013372 Homo sapi
12	825.8	90.8	1084	6	AX364851 Sequence
13	825.8	90.8	1106	6	AX364852 Sequence
14	701	77.1	993	10	AF457576 Mus muscu
15	701	77.1	993	10	AF543541 Mus muscu
16	541	59.5	1067	6	AR339362 Sequence
17	528	58.1	626	6	BD149917 Primer fo
18	528	58.1	626	6	AX869855 Sequence
19	474.8	52.2	1659	9	AK130203 Homo sapi

20	346.2	38.1	603	6	AX136449 Sequence
21	346.2	38.1	603	6	BD123689 Secretory
22	326.6	35.9	111084	9	AC006486 Homo sapi
23	325.6	35.8	1540	10	BC037043 Mus muscu
24	321	35.3	95663	9	AC010247 Homo sapi
25	301.4	33.2	303	6	AX431298 Sequence
26	280	30.8	226909	2	AC079490 Mus muscu
27	279.4	30.7	190669	2	AC120393 Homo sapi
28	250.2	27.5	167108	2	AC068283 Homo sapi
29	193.8	21.3	957	9	HS010973 Homo sapi
30	193.8	21.3	1139	6	BD227199 Protein f
31	193.8	21.3	1139	6	AX008886 Sequence
32	193.8	21.3	1142	6	AR477511 Sequence
33	193.8	21.3	1830	6	BD157757 Primer fo
34	193.8	21.3	1830	6	AX879289 Sequence
35	193.8	21.3	1830	9	AK022531 Homo sapi
36	193.8	21.3	2261	6	BD156359 Primer fo
37	193.8	21.3	2261	6	AX397469 Sequence
38	193.8	21.3	2261	6	AX876867 Sequence
39	193.8	21.3	2261	6	AK001497 Homo sapi
40	193.8	21.3	2624	5	BC084232 Xenopus 1
41	192.2	21.1	957	9	AF083236 Homo sapi
42	192.2	21.1	957	9	AF100341 Homo sapi
43	192.2	21.1	957	9	CR536556 Homo sapi
44	192.2	21.1	1045	6	AR140521 Sequence
45	192.2	21.1	1045	6	AR343032 Sequence

ALIGNMENTS

RESULT 1	BD093311	909 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093311				
DEFINITION	Apoptosis Relating Factor.				
ACCESSION	BD093311				
VERSION	BD093311.1	GI:22638899			
KEYWORDS	WO 0104300-A/1.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Miyoshi,S. and Sato,S.				
TITLE	1 (bases 1 to 909)				
JOURNAL	Apoptosis Relating Factor				
Patent:	WO 0104300-A 1 18-JAN-2001;				
HELIIX RESEARCH INSTITUTE,TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,SOUSUKE MIYOSHI,					
PI SUSUMU SATO					
OS Homo sapiens (human)					
PN WO 0104300-A/1					
PD 18-JAN-2001					
PF 06-JUL-2000 WO 2000JP004516					
PR 08-JUL-1999 JP 99P 194179,18-OCT-1999 US 60/159586					
TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,SOUSUKE MIYOSHI,					
PI SUSUMU SATO					
PC C12N15/12,C07K14/47,C12N5/10,C12N1/21,C12N1/19,C12N1/15,C12P21/					
02,					
PC C07K16/18,C12P21/08,G01N33/53,G01N33/577					
CC					
FT Key	Location/Qualifiers				
FT CDS	(1)..(909).				
FEATURES	Location/Qualifiers				
source	1..909				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	100.0%;	Score 909;	DB 6;	Length 909;	
Best Local Similarity	100.0%;	Pred. No. 1,2e-134;			
Matches 909;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	

QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGAGAGAGATGAGTCTTGACTAC 60
DB 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGAGAGAGATGAGTCTTGACTAC 60
QY 61 TAGGGAGATGATGCTGCTTCAACCGTATGTTAGAGTGTGTGGCGGGCAACTGACCGAGTGC 120
DB 61 TAGGGAGATGATGCTGCTTCAACCGTATGTTAGAGTGTGTGGCGGGCAACTGACCGAGTGC 120
QY 121 GAGCTGAGAGCTCTGAGCTTCTGCTGAGATGAGGCTCTGCGCCCGCGAGGCTTAGCC 180
DB 121 GAGCTGAGAGCTCTGAGCTTCTGCTGAGATGAGGCTCTGCGCCCGCGAGGCTTAGCC 180
QY 181 CGGGCCCGGACGCGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGAGCTGTGGCGAG 240
DB 181 CGGGCCCGGACGCGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGAGCTGTGGCGAG 240
QY 241 AGCAACCTGCGGCTGCTGCGGGCAACTCTGCGGTGCTGCGCCCGGCAAGCACTGCTGCGG 300
DB 241 AGCAACCTGCGGCTGCTGCGGGCAACTCTGCGGTGCTGCGCCCGGCAAGCACTGCTGCGG 300
QY 301 CACCTGCGCGACGAGCGCGCCGCGCAGTGTCTCAGAAAGCTATAGCTATGCACTCC 360
DB 301 CACCTGCGCGACGAGCGCGCCGCGCAGTGTCTCAGAAAGCTATAGCTATGCACTCC 360
QY 361 AGCTCTTCAAAAGAGACAGAGGAGTGTGCTGCGCTGCGGAGTCAAGCAAGTCTGCA 420
DB 361 AGCTCTTCAAAAGAGACAGAGGAGTGTGCTGCGCTGCGGAGTCAAGCAAGTCTGCA 420
QY 421 AATTCTCAGAGAGGTGAGTGGGAGACAGAGCTCTCCCGCCCAACGAGCGCGAGCGGAGT 480
DB 421 AATTCTCAGAGAGGTGAGTGGGAGACAGAGCTCTCCCGCCCAACGAGCGCGAGCGGAGT 480
QY 481 CGGGCCCGGACGAGTGTGTGCTCAGACGCGCGGAGAGAGGCGCCGAGCGGACCCAG 540
DB 481 CGGGCCCGGACGAGTGTGTGCTCAGACGCGCGGAGAGAGGCGCCGAGCGGACCCAG 540
QY 541 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGAATCTGATCTGAGCTC 600
DB 541 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGAATCTGATCTGAGCTC 600
QY 601 CGGGTTCAGAGAGTGTGCTGAGAGAGGCGAGGCGGCTTGGAGAGAGGCGGTGAGTCCCG 660
DB 601 CGGGTTCAGAGAGTGTGCTGAGAGAGGCGAGGCGGCTTGGAGAGAGGCGGTGAGTCCCG 660
QY 661 CGGCGCCGAGGCTGTGGCGCGGAGCTGAGAGTGTGTTGGAGGCAACGAGTGTCTGCG 720
DB 661 CGGCGCCGAGGCTGTGGCGCGGAGCTGAGAGTGTGTTGGAGGCAACGAGTGTCTGCG 720
QY 721 TCAAGGAGCTGTGGCTGTGTTGTGACATCAAGTCTCAAGCTCTCTATCTGAGC 780
DB 721 TCAAGGAGCTGTGGCTGTGTTGTGACATCAAGTCTCTATCTGAGC 780
QY 781 GCTTTCGAGGAGCACTGAGTGTGGCGCTGCTGAGAGGCGCTGCGGGCGGTCTTCTG 840
DB 781 GCTTTCGAGGAGCACTGAGTGTGGCGCTGCTGAGAGGCGCTGCGGGCGGTCTTCTG 840
QY 841 ACTGAGGCGCTGAGAGAGCTGTGTGGCGGAGGCTTCTGCGCTGAGTCAAGTGTGAT 900
DB 841 ACTGAGGCGCTGAGAGAGCTGTGTGGCGGAGGCTTCTGCGCTGAGTCAAGTGTGAT 900
QY 901 GAGGCTGAC 909
DB 901 GAGGCTGAC 909

RESULT 2
BD093312
LOCUS BD093312 1883 bp DNA linear PAT 27-AUG-2002
DEFINITION Apoptosis Relating Factor.
ACCESSION BD093312
VERSION BD093312.1 GI:22638900
*KEYWORDS WO 0104300-A/2.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1883)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Miyoshi,S. and Sato,S.
TITLE Apoptosis Relating Factor
JOURNAL Patent: WO 0104300-A 2 18-JAN-2001;
HELIX RESEARCH INSTITUTE,TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,
YURI KAWAI,SOUSUKE MIYOSHI,SUSUMU SATO
COMMENT OS Homo sapiens (human)
PN WO 0104300-A/2
PD 18-JAN-2001 WO 2000JP004516
PF 06-JUL-2000
PR 08-JUL-1999 JP 99P 194179,18-OCT-1999 US 60/159586 PI
TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,SOUSUKE
MIYOSHI,
PI SUSUMU SATO
PC C12N15/12,C07K14/47,C12N5/10,C12N1/21,C12N1/19,C12N1/15,C12P21/PC
02,C07K16/18,C12P21/08,G01N33/53,G01N33/577
CC Key Location/Qualifiers
FH Key (124)..(1101).
FT CDS location/Qualifiers
1..1883
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 909; DB 6; Length 1883;
Best Local Similarity 100.0%; Pred. No. 9,7e-135;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGAGAGAGATGAGTCTTGACTAC 60
DB 124 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGAGAGAGATGAGTCTTGACTAC 183
QY 61 TAGGGAGATGATGCTGCTTCAACCGTATGTTAGAGTGTGTGGCGGGCAACTGACCGAGTGC 120
DB 184 TAGGGAGATGATGCTGCTTCAACCGTATGTTAGAGTGTGTGGCGGGCAACTGACCGAGTGC 243
QY 121 GAGCTGAGAGCTCTGAGCTTCTGCTGAGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 180
DB 244 GAGCTGAGAGCTCTGAGCTTCTGCTGAGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 303
QY 244 GAGCTGAGAGCTCTGAGCTTCTGCTGAGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 303
DB 181 CGGGCCCGGACGCGCTTAGAGCTCTGCTGAGAGTGTGTTGGAGGCAACGAGTGTCTGCG 240
QY 304 CGGGCCCGGACGCGCTTAGAGCTCTGCTGAGAGTGTGTTGGAGGCAACGAGTGTCTGCG 363
DB 241 AGCAACCTGCGGCTGCTGCGGGCAACTCTGCGGTGCTGCGCCCGGCAAGCACTGCTGCGG 300
QY 364 AGCAACCTGCGGCTGCTGCGGGCAACTCTGCGGTGCTGCGCCCGGCAAGCACTGCTGCGG 423
DB 301 CACCTGCGCGACGAGCGCGCCGCGCAGTGTCTCAGAAAGCTATAGCTATGCACTCC 360
QY 424 CACCTGCGCGACGAGCGCGCCGCGCAGTGTCTCAGAAAGCTATAGCTATGCACTCC 483
DB 361 AGCTCTTCAAAAGAGACAGAGGAGTGTGCTGCGCTGCGGAGTCAAGCAAGTCTGCA 420
QY 484 AGCTCTTCAAAAGAGACAGAGGAGTGTGCTGCGCTGCGGAGTCAAGCAAGTCTGCA 543
DB 421 AATTCTCAGAGAGGTGAGTGGAGACAGGCTCTCCCGCCCAACGAGCGGAGCGGAGT 480
QY 544 AATTCTCAGAGAGGTGAGTGGAGACAGGCTCTCCCGCCCAACGAGCGGAGCGGAGT 603
DB 481 CGGGCCCGGACGAGTGTGTGCTCAGACGCGCGGAGAGAGGCGCCGAGCGGACCCAG 540
QY 604 CGGGCCCGGACGAGTGTGTGCTCAGACGCGCGGAGAGAGGCGCCGAGCGGACCCAG 663
DB 541 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGAATCTGATCTGAGCTC 600

Db 664 CAGAGTCAGAGCCCGCCGAGACCTTCTCTAAGGCAAGGACCTGTGACATCCGGCTC 723
Qy 601 CGGGTTGAGAGAGTAAGTCTGAGAGATGGGCGAGCTTTGAGAGAGGCGGTGGCATCCCGG 660
Db 724 CGGGTTGAGAGAGTAAGTCTGAGAGATGGGCGAGCTTTGAGAGAGGCGGTGGCATCCCGG 783
Qy 661 CGGCCCCAGAGCGCTGGCGCGGACAGTGGACCTGTTGGGAGCCGACCGCAGTGTGCGC 720
Db 784 CGGCCCCAGAGCGCTGGCGCGGACAGTGGACCTGTTGGGAGCCGACCGCAGTGTGCGC 843
Qy 721 TCAAGGAGCTGGGCTCTGTGTGTTGTGACATCAAGTCTCTAGAGCTTCTTATCTGAC 780
Db 844 TCAAGGAGCTGGGCTCTGTGTGTTGTGACATCAAGTCTCTAGAGCTTCTTATCTGAC 903
Qy 781 GCGTTCTGGGCGGAGTACTAGTGGCGGCTCTGCTGAGGCGCTTGGGCGGTGTTCTG 840
Db 904 GCGTTCTGGGCGGAGTACTAGTGGCGGCTCTGCTGAGGCGCTTGGGCGGTGTTCTG 963
Qy 841 ACTGAGGCGCTGGAGAGGCTGTGGGCGGAGGCGTGTTCGCTGCTGTCAGTGTGGAT 900
Db 964 ACTGAGGCGCTGGAGAGGCTGTGGGCGGAGGCGTGTTCGCTGCTGTCAGTGTGGAT 1023
Qy 901 GAGGCTGAC 909
Db 1024 GAGGCTGAC 1032

RESULT 3
AK075328 1883 bp mRNA linear PRI 03-SEP-2002
LOCUS AK075328
DEFINITION Homo sapiens cDNA PSEC0004 f1s, clone NT2RM1000558, weakly similar to Homo sapiens death effector domain-containing testicular molecule RNA.
ACCESSION AK075328
VERSION AK075328.1 GI:22761346
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Hayashi,K., Ishii,S., Saito,K., Yamamoto,J., Wakamatsu,A., Nagai,T., Nakamura,Y., Negahari,K., Sugano,S. and Isegai,T.
TITLE HRI human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 1883)
Isegai,T. and Yamamoto,J.
COMMENT Direct Submission
Submitted (20-MAR-2002) Takao Isegai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass sequencing, clone selection and full insert sequencing; Helix Research Institute (supported by Japan Key Technology Center etc.); cDNA library construction; Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES
Source Location/Qualifiers
1.1883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM1000558"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RM1"
/note="cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells"
124.1104
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11551.1"

/db_xref="GI:22761347"
/translation="MALSGSTPAPCWEDECLDYGLSLHMEPVVGGQLTBCLEL
LAFLLDEPAGAGGLARSGLELLERGGQGESNLRLQLLVLRHDLPH,
ARKRRVSVSPERYSYGTSSSKRTGECRRRQSSSANSQQQWETGSPYTKQRS
RGRSGGARRRRGATPAPOOSEPAPRSEGGKTCIDIRLRVAREYCEHPALEQVA
SRPQALAROLDVNGATAVLRSDIGSVICDIKFSLSYDAEMKGYLSGALLQLAR
GVFLTEALREAVRGAVRLVSVDEADYEAGRRRLLMEEBGRPRTEAS"

ORIGIN
Query Match 100.0%; Score 909; DB 9; Length 1883;
Best Local Similarity 100.0%; Pred. No. 9 7e-135;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGCGCTATCCGGGTCGACCCCGGCGCCGCTGCTGGAGAGAGATGAGTGTGACTTAC 60
Db 124 ATGGCGCTATCCGGGTCGACCCCGGCGCCGCTGCTGGAGAGAGATGAGTGTGACTTAC 183
Qy 61 TACGGGATGCTGTGTGCTTACCCGTAATGTTGAGAGTGGTGGCGGGCAACTGACGAGTGC 120
Db 184 TACGGGATGCTGTGTGCTTACCCGTAATGTTGAGAGTGGTGGCGGGCAACTGACGAGTGC 243
Qy 121 GAGCTGAGAGCTCTGAGCTTCTGTGATGAGGCTCTGAGCGCGCGCGAGGCTTAGCC 180
Db 244 GAGCTGAGAGCTCTGAGCTTCTGTGATGAGGCTCTGAGCGCGCGCGAGGCTTAGCC 303
Qy 181 CGGCGCCGAGCGGCTTAGAGCTTCTGTGAGCTGAGCGCGCGGCGCACTGCGCGAG 240
Db 304 CGGCGCCGAGCGGCTTAGAGCTTCTGTGAGCTGAGCGCGCGGCGCACTGCGCGAG 363
Qy 241 AGCAACTGCGGCGCTGTGGGCAACTCTGTGCGTGTGCTGCGCGCGCACTGCTGCGG 300
Db 364 AGCAACTGCGGCGCTGTGGGCAACTCTGTGCGTGTGCTGCGCGCGCACTGCTGCGG 423
Qy 301 CACCTGGCGGCAAGCGGCGCGCGCGAGTGTCTCAGAGGCTATAGTATGAGACCTCC 360
Db 424 CACCTGGCGGCAAGCGGCGCGCGAGTGTCTCAGAGGCTATAGTATGAGACCTCC 483
Qy 361 AGCTCTTCAAAAGAGAGAGAGGAGTGTGCTGCGCGCTGCGCGAGTGTGCTGCA 420
Db 484 AGCTCTTCAAAAGAGAGAGAGGAGTGTGCTGCGCGCTGCGCGAGTGTGCTGCA 543
Qy 421 AATTCTCAGCAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGCGAGGAGT 480
Db 544 AATTCTCAGCAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGCGAGGAGT 603
Qy 481 CGGGCGGCGGCAATGTGTGTGTCAGACGCGGCGGAGAGGCGGCCACGCGCACCCAG 540
Db 604 CGGGCGGCGGCAATGTGTGTGTCAGACGCGGCGGAGAGGCGGCCACGCGCACCCAG 663
Qy 541 CAGCAGTCAGAGCCCGGCAAGCTTCTCTGAAAGCAAAAGTGAACCTGTGACATCCGGCTC 600
Db 664 CAGCAGTCAGAGCCCGGCAAGCTTCTCTGAAAGCAAAAGTGAACCTGTGACATCCGGCTC 723
Qy 601 CGGGTTGAGAGAGTAAGTCTGAGAGATGGGCGAGCTTTGAGAGAGGCGGTGGCATCCCGG 660
Db 724 CGGGTTGAGAGAGTAAGTCTGAGAGATGGGCGAGCTTTGAGAGAGGCGGTGGCATCCCGG 783
Qy 661 CGGCCCCAGAGCGCTGGCGCGGACAGTGGACCTGTTGGGAGCCGACCGCAGTGTGCGC 720
Db 784 CGGCCCCAGAGCGCTGGCGCGGACAGTGGACCTGTTGGGAGCCGACCGCAGTGTGCGC 843
Qy 721 TCAAGGAGCTGGGCTCTGTGTGTTGTGACATCAAGTCTCTAGAGCTTCTTATCTGAC 780
Db 844 TCAAGGAGCTGGGCTCTGTGTGTTGTGACATCAAGTCTCTAGAGCTTCTTATCTGAC 903
Qy 781 GCGTTCTGGGCGGAGTACTAGTGGCGGCTCTGCTGAGGCGCTTGGGCGGTGTTCTG 840
Db 904 GCGTTCTGGGCGGAGTACTAGTGGCGGCTCTGCTGAGGCGCTTGGGCGGTGTTCTG 963
Qy 841 ACTGAGGCGCTGGAGAGGCTGTGGGCGGAGGCGTGTTCGCTGCTGTCAGTGTGGAT 900
Db 964 ACTGAGGCGCTGGAGAGGCTGTGGGCGGAGGCGTGTTCGCTGCTGTCAGTGTGGAT 1023

Qy	901	GAGCTGAC	909
Db	1024	GAGCTGAC	1032

RESULT 4
CQ719642

LOCUS	CQ719642	981 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	Sequence 5576 from Patent WO02068579.				

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.	Kits, such as nucleic acid arrays, comprising a majority of

JOURNAL	Patent: WO 02068579-A 5576 06-SEP-2002
PE Corporation (NY) (US)	
FEATURES	Location/Qualifiers
source	1. .981

ORIGIN

Query Match	99.8%	Score 907.4	DB 6	Length 981
Best Local Similarity	99.9%	Pred. No. 2.1e-134		
Matches 908	Conservative	0	Mismatches 1	Indels 0
				Gaps 0

QY	1	ATGCGGCTATCCGGGTGCAACCCCGGCCCGGTCTGGGAGAGAGATGATGCTCGACCTAC	60
Db	1	ATGGGGCTATCCGGGTGCAACCCCGGCCCGGTCTGGGAGAGAGATGATGCTCGACCTAC	60
QY	61	TACGGGATGCTGTCCGCTTCAACCGTATGTTCCAGAGTGTGGGCGGGCAACTGACCCGATGC	120
Db	61	TACGGGATGCTGTCCGCTTCAACCGTATGTTCCAGAGTGTGGGCGGGCAACTGACCCGATGC	120
QY	121	GAGCTGAGAGCTCTGAGCCCTTTCTGCTGATGAGGCTCTGGCGCGCCGAGGCTTAAACC	180
Db	121	GAGCTGAGAGCTCTGAGCCCTTTCTGCTGATGAGGCTCTGGCGCGCCGAGGCTTAAACC	180
QY	181	CGGGCCCCCAGCGGGCTAGAGCTCTCGCTGGAGCTGGAGACGCGCGGGGACATGGCCCGAG	240
Db	181	CGGGCCCCCAGCGGGCTAGAGCTCTCGCTGGAGCTGGAGACGCGCGGGGACATGGCCCGAG	240
QY	241	AGCAACTGCGGCTCTGTGGGGCAACTCTGCGCGTGTGAGCCGCGCAGCACTGTGCGG	300
Db	241	AGCAACTGCGGCTCTGTGGGGCAACTCTGCGCGTGTGAGCCGCGCAGCACTGTGCGG	300
QY	301	CACCTGCGCGCAGCGGCGCGGCCAGTGTTCCAGAAAGCTTATGCTATGGACCTCC	360
Db	301	CACCTGCGCGCAGCGGCGCGGCCAGTGTTCCAGAAAGCTTATGCTATGGACCTCC	360
QY	361	AGCTTTTCAANAAGACAGAGGGTATGTCGCGCGCGTCCGAGATCAAGCACTTCTGCA	420
Db	361	AGCTTTTCAANAAGACAGAGGGTATGTCGCGCGCGTCCGAGATCAAGCACTTCTGCA	420
QY	421	AATTCTCAGCAGGCTCAGTGGGAGACAGGCTCCCCCAACCAAGAGGCGACCGGCGAGT	480
Db	421	AATTCTCAGCAGGCTCAGTGGGAGACAGGCTCCCCCAACCAAGAGGCGACCGGCGAGT	480
QY	481	CGGGCCCGGCCAGTGTGTGTCCAGACGGCGCGGAGAGGGGCCCCAGCGCACCCCAAG	540
Db	481	CGGGCCCGGCCAGTGTGTGTCCAGACGGCGCGGAGAGGGGCCCCAGCGCACCCCAAG	540
QY	541	CAGCAGTCAGAGCCCGCCAGACCTTCTCTTGAAGGCAAAAGTACCTGTGATCATCCGGGCTC	600
Db	541	CAGCAGTCAGAGCCCGCCAGACCTTCTCTTGAAGGCAAAAGTACCTGTGATCATCCGGGCTC	600

QY	CGGGTTGAGACAGTACTGTGACATAGGCGCAGCTTGGAGCAGAGGGGTGACATCCGG	601
Db	CGGGTTGAGCAGAGTACTGTGACATAGGCGCAGCTTGGAGCAGAGGGGTGACATCCGG	601
QY	CGGCCCAAGCGCTGAGCGCGCAGCTGGACGTGTTTGGGCAGGCCACCGAGTGTGCGC	661
Db	CGGCCCAAGCGCTGTGCGCGGCACTGGAAGTGTGTTTGGGCAGGCCACCGAGTGTGCGC	661
QY	TCAAGGAGCCTGGGCTCTGTGGTTTGTGACATCAAGTTCTCAGAGCTCTCCTATCTGAC	721
Db	TCAAGGAGCCTGGGCTCTGTGGTTTGTGACATCAAGTTCTCAGAGCTCTCCTATCTGAC	721
QY	GCCTTTCTGGGGCGACTACCTGAGTGGCGCCCTGCTGACAGGCGCTTCGGGGGCGTGTCTTG	781
Db	GCCTTTCTGGGGCGACTACCTGAGTGGCGCCCTGCTGACAGGCGCTTCGGGGGCGTGTCTTG	781
QY	ACTAGAGCCCTGCGAGAGCGCTGTGGCGCGGAGGCGCTGTTGCGCTGAGTGTGAT	841
Db	ACTAGAGCCCTGCGAGAGCGCTGTGGCGCGGAGGCGCTGTTGCGCTGAGTGTGAT	841
QY	GAGGCTGAC 909	901
Db	GAGGCTGAC 909	901

RESULT 5
APA43591

SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE

AUTHORS	KOTLI, W., Stenner-Jewett, F., PAWLOWSKI, K., GADZIK, A. and KEEA, J.C.
TITLE	Identification and characterization of DEDD2, a death effector domain-containing protein

REFERENCE 2 (bases 1 to 981)
AUTHORS Roth, W. and Reed, J. C.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) Burnham Institute, 10901 N. Torrey Pines

FEATURES	Location/Qualifiers
source	1. .981

gene	1. .981
CDS	/gene="DEDD2" 1. .981

```

/codon_start=1
/product="death effector domain-containing DNA-binding
protein 2"

```

protein_id="PAL48220.1"
/db_xref="GI:17933267"
/translation="MALSGSRPAPCMEDECLDYYGMLSLHMEVVGGLTDEL
LAFILDEAPGAGGLARARSGLELLLEERGGQDSENRLTGLGLRLVLAHLDLPHIL
AKRRRVPSPERSVYGTSSSKRTPEGSPRRROSSSNANSGQWETSPTPKRRR
RGRRSGGARRRRRGAAPAPAPQOQSEAPSPSGKTCIDIRLVRAVCHSGALGEGV
SRRLQALARCILDVGAQTAVLRSDRLSDVDCIFSELSYDAWAGVLTGALLQALML
GVFTLEALREAVGEAVRLVSVDEAVYVGGDRLFLMEEGGRRPTEAS"

ORIGIN

Query Match 99.8%; Score 907.4; DB 9; Length 981;


```

Db      847 TCAAGGACCTGGGCTCTGTGTTTGATCATCAAGTCTCAAGAGCTCTCTATCTGGAC
Qy      781 GCCTTGGGGGCACTACTAGTGGCGCCCTGCTGAGAGCCCTGCGGGGCGTGTTCCTG
Db      907 GCCTTGGGGGCACTACTAGTGGCGCCCTGCTGAGAGCCCTGCGGGGCGTGTTCCTG
Qy      841 ACTGAGGCGCTGCGAGAGGCTGTGGCCGCGAGAGCTGTTCCTGCTGTCAAGTGGAT
Db      967 ACTGAGGCGCTGCGAGAGGCTGTGTGGCCGCGAGAGCTGTTCCTGCTGTCAAGTGGAT
Qy      901 GAGGCTGAC 909
Db      1027 GAGGCTGAC 1035

RESULT 7
LOCUS   AX431308      1924 bp      DNA      linear      PAT 28-JUN-2002
DEFINITION
ACCESSION AX431308
VERSION   AX431308.1 GI:21656177
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS   Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C.,
          Roth, W. and Stenier-Liemen, F.
          Novel death domain proteins
          Patent: WO 0240680-A 17 23-MAY-2002;
          BURNHAM INST (US)
FEATURES
    source
        1. 1924
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"
                91..>1044
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAD37280.1"
                /db_xref="GI:21656178"
                /translation="MALSGSTPAPCWEDECDLYGMLSLHMEFVVGQLTCELEL
                /LAFLLDEAPAGGLARARSGLELLELRGQCDSENLRLGLQLLVLARHDLPHL
                /ARKRRVSPERYSYGTSSSKRTGSSRRSSSANSQCGWETGSPTRKORS
                /RGRPSGGARRRRGAPAPAPPOQSEPARPSEGGKTCIDIRLVRAREYCEHGPALBOGVA
                /SRPQALAROLDVPGCATATLRSRDLSVVICDIKESPLSYLIDAFWMDYLSGALLQALR
                /GVFLTEALREAVGREAVRLVSVDEADYEAGRRRLIMEERG"

ORIGIN
Query Match      99.8%; Score 907.4; DB 6; Length 1924;
Best Local Similarity 99.9%; Pred. No. 1.7e-134;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGCGCGTATCCGGGTCGACCCCGGCCCGCTGCTGGAGAGAGATGATGCTGAGCTAC 60
Db      91 ATGCGCGTATCCGGGTCGACCCCGGCCCGCTGCTGGAGAGAGATGATGCTGAGCTAC 150
Qy      61 TACGGGATGCTGTGCTTCAACCGATGTTTCAGAGTGTGTGGCGGGGCAACTGACCGAGTGC 120
Db      151 TACGGGATGCTGTGCTTCAACCGATGTTTCAGAGTGTGTGGCGGGGCAACTGACCGAGTGC 210
Qy      121 GAGCTGAGGCTCTGGGCTTTCTGCTGGATGAGGCTCTGGCGCGCGCGAGGCTTACGCC 180
Db      211 GAGCTGAGGCTCTGGGCTTTCTGCTGGATGAGGCTCTGGCGCGCGCGAGGCTTACGCC 270
Qy      181 CGGGCCCGCAGCGGCTTACAGAGCTCTGCTGAGAGTGAAGCGCGCGGCACTGCGCGAG 240
Db      271 CGGGCCCGCAGCGGCTTACAGAGCTCTGCTGAGAGTGAAGCGCGCGGCACTGCGCGAG 330
Qy      241 AGCAACTGCGGCTGTGGGCAACTCTGCGCGCTGCTGCGCGCGCGCAAGCTGCTGCCG 300

```

```

Db      331 AGCAACTGCGGCTGTGGGCAACTCTGCGCGCTGCTGCGCGCGCGCAAGCTGCTGCCG 390
Qy      301 CACTTGGCGGCAAGCGGCGCGCGCAAGTGTCTCCAGAAAGCTATAGTACACTTCC 360
Db      391 CACTTGGCGGCAAGCGGCGCGCGCAAGTGTCTCCAGAAAGCTATAGTACACTTCC 450
Qy      361 AGCTTTCAAAAGAGACAGAGGTTACTGTCCTGCGCGCGCGCGCAAGTCAAGCACTTCCGA 420
Db      451 AGCTTTCAAAAGAGACAGAGGTTACTGTCCTGCGCGCGCGCGCAAGTCAAGCACTTCCGA 510
Qy      421 AATTCTCAGCAGGGTCAATGAGAGACAGGCTCCCGCCCAACCAAGCGGCGCGCGAGT 480
Db      511 AATTCTCAGCAGGGTCAATGAGAGACAGGCTCCCGCCCAACCAAGCGGCGCGAGT 570
Qy      481 CGGGCCCGGCGCAAGTGTGTGTCAGACCGCGCGGAGAGGGGCCCGCAGCCGACCCAG 540
Db      571 CGGGCCCGGCGCAAGTGTGTGTCAGACCGCGCGGAGAGGGGCCCGCAGCCGACCCAG 630
Qy      541 CAGCAGTCAAGAGCCCGCCAGACCTTCTCTGGAAGGCAAGTGAACCTGTGACATCCGGCTC 600
Db      631 CAGCAGTCAAGAGCCCGCCAGACCTTCTCTGGAAGGCAAGTGAACCTGTGACATCCGGCTC 690
Qy      601 CGGATTCAAGCAGAGTACTGCGAGCATGAGGCGCAGCCTTGGAGAGGAGCGTGATCCCG 660
Db      691 CGGATTCAAGCAGAGTACTGCGAGCATGAGGCGCAGCCTTGGAGAGGAGCGTGATCCCG 750
Qy      661 CGGCCCCAGGCGCTGCGCGCGCGCAGCTGAGCATGTGTGGGCGAGCCACCGAGTCTGCGC 720
Db      751 CGGCCCCAGGCGCTGCGCGCGCGCAGCTGAGCATGTGTGGGCGAGCCACCGAGTCTGCGC 810
Qy      721 TCAAGGACCTGGGCTCTGTGTTTGTGACATCAAGTCTCAAGCTCTCTATCTGGAC 780
Db      811 TCAAGGACCTGGGCTCTGTGTTTGTGACATCAAGTCTCAAGCTCTCTATCTGGAC 870
Qy      781 GCCTTGGGGGCACTACTAGTGGCGCCCTGCTGAGAGCCCTGCGGGGCGTGTTCCTG 840
Db      871 GCCTTGGGGGCACTACTAGTGGCGCCCTGCTGAGAGCCCTGCGGGGCGTGTTCCTG 930
Qy      841 ACTGAGGCGCTGCGAGAGGCTGTGGCCGCGAGAGCTGTTCCTGCTGTCAAGTGGAT 900
Db      931 ACTGAGGCGCTGCGAGAGGCTGTGTGGCCGCGAGAGCTGTTCCTGCTGTCAAGTGGAT 990
Qy      901 GAGGCTGAC 909
Db      991 GAGGCTGAC 999

RESULT 8
LOCUS   BC027930      1951 bp      mRNA      linear      PRI 29-JUN-2004
DEFINITION Homo sapiens death effector domain containing 2, mRNA (cDNA clone
ACCESSION BC027930
VERSION   BC027930.1 GI:20379818
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1951)
AUTHORS   Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
          Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
          Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, X., Gibbs, R.A.,
          Vittalion, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

```


Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Boufford, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimmwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalek, U., Smalls, D. E., Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1951)

Strausberg, R.
Direct Submission
Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapdb-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Ahter, N., Ayale, K., Beckstrom-Sternberg, S. M., Benjamin, B., Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N. L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q. L., Mastello, C., Maekert, B., Mastrian, S. D., McCloskey, J. C., McDowell, J., Pearson, R., Stantrilop, S., Thomas, P. J., Touchman, J. W., Tourgon, C., Vogt, J. L., Walker, M. A., Weherby, K. D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19921049.

Location/Qualifiers

1. 1951

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:35529 IMAGE:5199772"

/tissue_type="Brain, adult, 6 pooled whole brains"

/clone_id="NIH_MGC_114"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 1951

/gene="DEDD2"

/note="synonym: FLAME-3"

/db_xref="LocusID:162989"

69. 1049

/gene="DEDD2"

/codon_start=1

/product="death effector domain-containing DNA binding protein 2"

/protein_id="AAH27930.1"

/db_xref="GI:20379819"

/db_xref="LocusID:162989"

/translation="MALSGSTPAPCWEDECLDYGMLSHRMEVVGQUTCELELAFILDEAPAGAGLARARSGLELLLEIRGQCDSENRLQLVLVAHDLPLHAKRRRVSPEPVYGTSSSKRTGSGLEERSSSNSSNQOGWETGSPETKQRRSRPSGAGARRRRGAPAPQOQSEPARSPSEKTYCDIIRLRAVEYCEHGALEQVARSPOALAROLVDGOATAVARSRDLSVDCIKFSELSYDAFMGDIYLSGALLQALRGVFLTEALREAVRGREAVRLVSVDEADYEARRRLLMEEGGRPTBAS"

ORIGIN

Query Match 99.5%; Score 904.2; DB 9; Length 1951;

Best Local Similarity 99.7%; Pred. No. 5,6e-134;

	Matches	906; Conservative	0; Mismatches	3; Indels	0; Gaps	0;
Qy	1	ATGCGCTATCCGGGTGACCCCGCGCGTGTGAGAGAGATGACTGACTAC	60			
Db	69	ATGCGCTATCCGGGTGACCCCGCGCGTGTGAGAGAGATGACTGACTAC	128			
Qy	61	TACGGATGCTGTGCTTCAACCGTATGAGGTGAGCGGCACTGACCGAGTGC	120			
Db	129	TACGGATGCTGTGCTTCAACCGTATGAGGTGAGCGGCACTGACCGAGTGC	188			
Qy	121	GAGTGAAGCTCTCGGCTTTCTGCTGATGAGGCTCTCGGCGCGGAGGCTTGCC	180			
Db	189	GAGTGAAGCTCTCGGCTTTCTGCTGATGAGGCTCTCGGCGCGGAGGCTTGCC	248			
Qy	181	CGGCGCGGAGCGGCTGAGAGCTCTGCTGAGTGAAGCGCGCGGAGTGAAG	240			
Db	249	CGGCGCGGAGCGGCTGAGAGCTCTGCTGAGTGAAGCGCGCGGAGTGAAG	308			
Qy	241	AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCGTGTGCGCGCAAGACTGCGCG	300			
Db	309	AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCGTGTGCGCGCAAGACTGCGCG	368			
Qy	301	CACCTGGCGGCAAGCGGCGCGGCACTGCTTCAAGAGCTATAGTACACTCC	360			
Db	369	CACCTGGCGGCAAGCGGCGCGGCACTGCTTCAAGAGCTATAGTACACTCC	428			
Qy	361	AGCTCTCAAG	420			
Db	429	AGCTCTCAAG	488			
Qy	421	AATTCAGAGAGGTCACTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	480			
Db	489	AATTCAGAGAGGTCACTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	548			
Qy	481	CGGCGCGGCGGCACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	540			
Db	549	CGGCGCGGCGGCACTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	608			
Qy	541	CAGAGTCAGAGCGCGGCAAGCTTCTGTAAGGCAAGTGAATGATGATCGGCTC	600			
Db	609	CAGAGTCAGAGCGCGGCAAGCTTCTGTAAGGCAAGTGAATGATGATCGGCTC	668			
Qy	601	CGGCTTCAGAGAGTACTGAGCATGAGGCAAGCTTTCAGAGAGAGAGAGAGAG	660			
Db	669	CGGCTTCAGAGAGTACTGAGCATGAGGCAAGCTTTCAGAGAGAGAGAGAGAG	728			
Qy	661	CGGCGCGGCGGCTGCGCGGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720			
Db	729	CGGCGCGGCGGCTGCGCGGCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	788			
Qy	721	TCAAGGAGCTGAGGCTGTGCTTGTGACATCAAGTTCAGAGCTCTCTATCTGAG	780			
Db	789	TCAAGGAGCTGAGGCTGTGCTTGTGACATCAAGTTCAGAGCTCTCTATCTGAG	848			
Qy	781	GCTTTCAGGAGGCTGAGTGTGAGGCGCGCTGTCGAGGCGCGGCGTGTCTCG	840			
Db	849	GCTTTCAGGAGGCTGAGTGTGAGGCGCGCTGTCGAGGCGCGGCGTGTCTCG	908			
Qy	841	ACTGAGCGCTGCGAGAGGCTGTGAGGCGCGGAGGCTTGTGCTGCTGCTGAT	900			
Db	909	ACTGAGCGCTGCGAGAGGCTGTGAGGCGCGGAGGCTTGTGCTGCTGCTGAT	968			
Qy	901	GAGGCTGAC 909				
Db	969	GAGGCTGAC 977				
RESULT 9						
LOCUS	AF457575	981 bp	mRNA	linear	PRI 10-APR-2002	
DEFINITION	Homo sapiens death effector domain-containing protein FLAME-3					
ACCESSION	AF457575					

VERSION AF457575.1 GI:20126793
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 981)
Zhan,Y., Hegde,R., Srinivasula,S.M., Fernandes-Alnemri,T. and
Alnemri,E.S.
TITLE Death effector domain-containing proteins DEDD and FLAME-3 form
nuclear complexes with the TIF1C102 subunit of human transcription
factor IIIC
JOURNAL Cell Death Differ. 9 (4), 439-447 (2002)
MEDLINE 21961615
PUBMED 11965497
REFERENCE 2 (bases 1 to 981)
AUTHORS Alnemri,E.S.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2001) Microbiology and Immunology, Thomas
Jefferson University, 233 S. 10th Street, Philadelphia, PA 19107,
USA
FEATURES
source Location/Qualifiers
1..981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
1..981
/gene="FLAME-3"
1..981
/gene="FLAME-3"
/codon_start=1
/product="death effector domain-containing protein
FLAME-3"
/protein_id="AA010835.1"
/db_xref="GI:20126794"
/translation="MALSGSTPAPCWEDEBCELDYGMLSINRFVVGQLTRECELE
LAFLEDPAPGAGGLARSGLELLELEKRGCDSSNRLQLQLRLARHDLPHL
ARRRRPVSRRYSYGTSSSKRTESCRARRQSSANSOQOMETGSPTRORS
RGRPSGARRRRGAAPQOQSEPARPSEGVCDIRLVAAEYREHGPALQGV
SRPDLARLDVFGOATVLRSDIDGKVFSELSTLDLAFMDVLYALQLR
GVFLTEALREAVREAVRLVSDVDYENGRRRLIMEBEGRRPTAS"

Query Match 98.9%; Score 899.4; DB 9; Length 981;
Best Local Similarity 99.3%; Pred. No. 3.9e-133;
Matches 903; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

ORIGIN
QY 1 ATGGCGCTATCCGGGCTGACCCCGCCCGTCTGGAGAGAGATGCTCTGACTAC 60
DB 1 ATGGCGCTATCCGGGCTGACCCCGCCCGTCTGGAGAGAGATGCTCTGACTAC 60
QY 61 TACGGGATGCTGCGCTTACCGGATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
DB 61 TACGGGATGCTGCGCTTACCGGATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
QY 121 GAGCTGAGAGCTCTGGCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAGCC 180
DB 121 GAGCTGAGAGCTCTGGCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAGCC 180
QY 121 GAGCTGAGAGCTCTGGCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAGCC 180
DB 121 GAGCTGAGAGCTCTGGCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAGCC 180
QY 181 CGGGCGCGCGCGCGCTTAGAGCTCTGCTGAGTGGAGCGCGCGCGGCACTGCGGCGAG 240
DB 181 CGGGCGCGCGCGCGCTTAGAGCTCTGCTGAGTGGAGCGCGCGCGGCACTGCGGCGAG 240
QY 241 AGCAACCTGGGCTGCTGGGCGCACTCTGCGGCTGCGCGCGCGCGCGAGCTGCGCG 300
DB 241 AGCAACCTGGGCTGCTGGGCGCACTCTGCGGCTGCGCGCGCGCGCGAGCTGCGCG 300
QY 301 CACTTGGCGGCAAGCGCGCGCGCGCACTGCTCCAGAACGCTATGAGTACCTCC 360
DB 301 CACTTGGCGGCAAGCGCGCGCGCGCACTGCTCCAGAACGCTATGAGTACCTCC 360

QY 361 AGCTTTCAAAGAGCAGAGGGTAGCTGCGCGCGCGCGCGAGTCAAGCACTTTCGA 420
DB 361 AGCTTTCAAAGAGCAGAGGGTAGCTGCGCGCGCGCGCGAGTCAAGCACTTTCGA 420
QY 421 AATTTCACAGAGGGTAGTGGAGACAGGCTCCCCCAACCAAGGCGGAGCGGAGT 480
DB 421 AATTTCACAGAGGGTAGTGGAGACAGGCTCCCCCAACCAAGGCGGAGCGGAGT 480
QY 481 CGGGGCGCGCGCGAGTGGTGGTGCAGACGCGCGCGGAGAGGGGCGCCAGCGCAG 540
DB 481 CGGGGCGCGCGCGAGTGGTGGTGCAGACGCGCGCGGAGAGGGGCGCCAGCGCAG 540
QY 541 CAGCACTCAGAGCGCGCGCGCGAGCTTCTCTGAAAGCAAGTCACTGTACATCCGGCTC 600
DB 541 CAGCACTCAGAGCGCGCGCGCGAGCTTCTCTGAAAGCAAGTCACTGTACATCCGGCTC 600
QY 601 CGGGTTCAGACAGAGTACTGCGAGCATGAGGCGCAGCTTGGAGAGGAGCGTGGCATCCGG 660
DB 601 CGGGTTCAGACAGAGTACTGCGAGCATGAGGCGCAGCTTGGAGAGGAGCGTGGCATCCGG 660
QY 661 CGGCCCCAGGCGCTGGCGGCGGAGCTGACGTTGTTGGACAGGCCACCGCATGCTGCCG 720
DB 661 CGGCCCCAGGCGCTGGCGGCGGAGCTGACGTTGTTGGACAGGCCACCGCATGCTGCCG 720
QY 721 TCAGAGGACCTGGGCTCTGTTGTGATCAATCAAGTTCAGAGCTCTTATCTGGAC 780
DB 721 TCAGAGGACCTGGGCTCTGTTGTGATCAATCAAGTTCAGAGCTCTTATCTGGAC 780
QY 781 GCCTTCGGGCGAGTACTGAGTGGCGGCGGAGGCTGTTCCCTGCTGATGATGAT 840
DB 781 GCCTTCGGGCGAGTACTGAGTGGCGGCGGAGGCTGTTCCCTGCTGATGATGAT 840
QY 841 ACTGAGGCGCTGCGAGAGGCTGTTGGCGGAGGCTGTTCCCTGCTGATGATGAT 900
DB 841 ACTGAGGCGCTGCGAGAGGCTGTTGGCGGAGGCTGTTCCCTGCTGATGATGAT 900
QY 901 GAGGCTGAC 909
DB 901 GAGGCTGAC 909
RESULT 10
AY125488
LOCUS 2012 bp mRNA linear PRI 18-SEP-2002
DEFINITION Homo sapiens death effector domain-containing DNA-binding 2
ACCESSION AY125488
VERSION AY125488.1 GI:22475163
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 2012)
Lee,J.C., Schickling,O., Stegh,A.H., Oshima,R.G., Dinedale,D.,
Cohen,G.W., and Peter,M.E.
TITLE DEDD regulates degradation of intermediate filaments during
apoptosis
JOURNAL J. Cell Biol. 158 (6), 1051-1066 (2002)
MEDLINE 2220288
PUBMED 12235123
REFERENCE 2 (bases 1 to 2012)
AUTHORS Peter,M.E.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2002) Ben May Cancer Institute, University of
Chicago, 924 E. 57th Street, R112, Chicago, IL 60637, USA
FEATURES
source Location/Qualifiers
1..2012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2012
/gene="DEDD2"

CDS	84..1049	/gene="DED2"
	/codon_start=1	/product="death effector domain-containing DNA-binding 2 protein"
	/protein_id="AA095240.1"	/db_xref="GI:22475164"
	/translation="MALSGSTPAPCWEDECLDYVMISLHMFVVGQUTJCELEL LAFILDEPAGAGILARASGLLELLEERGQCDENLRLQLRLATLADLPHL ARKRRPVPSPRYVGTSSSKRTGSSRRSSSSANSQSPPTRRORSGRPS GGARRRRGAPAAPQOSEPARSPSEKVTCDILRYAAYEKDILAEQVARRPQ ALARQDVFGATATVLRSDISGVCDIKFSELSYLDPMFDYISGALLOLRGVFLT EALREAVREARLVSVDEADYENGRRLLMEEBGRPTLAS"	
ORIGIN	95.2%; Score 865.8; DB 9; Length 1012;	
Query Match	Best Local Similarity 98.1%; Pred. No. 6.8e-128;	
Matches	892; Conservative 0; Mismatches 2; Indels 15; Gaps 1;	
Qy	1 ATGGCGCTATCGGGTGCACCCCGGCCCGCTGCTGGAGAGATGCTGCTGACTAC	60
Db	84 ATGGCGCTATCGGGTGCACCCCGGCCCGCTGCTGGAGAGATGCTGCTGACTAC	143
Qy	61 TACGGGATGCTGCTGCTTCAACCGTATGTTCCAGTGTGGGCGGCAACTGACGAGTGC	120
Db	144 TACGGGATGCTGCTGCTTCAACCGTATGTTCCAGTGTGGGCGGCAACTGACGAGTGC	203
Qy	121 GAGCTGGAGCTCTGGGCTTCTGCTGGATGAGGCTCTGGGCGGCGGAGGCTTACC	180
Db	204 GAGCTGGAGCTCTGGGCTTCTGCTGGATGAGGCTCTGGGCGGCGGAGGCTTACC	263
Qy	181 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGTGAAGCGCGCGGCGAGTGCAGCAG	240
Db	264 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGTGAAGCGCGCGGCGAGTGCAGCAG	323
Qy	241 AGCAACTGCGGCTGCTGGGCACTCTGCGCGTGTGCGCCGCAAGACTGTGTCG	300
Db	324 AGCAACTGCGGCTGCTGGGCACTCTGCGCGTGTGCGCCGCAAGACTGTGTCG	383
Qy	301 CACCTGGGCGGCAAGCGGCGCGGCACTGCTCCAGAACGCTATAGTATGAGCACTCC	360
Db	384 CACCTGGGCGGCAAGCGGCGCGGCACTGCTCCAGAACGCTATAGTATGAGCACTCC	443
Qy	361 AGCTCTTCAAGAGAGACAGAGGTAGTCCGCTGCGCGAGTCAAGAGATTCTCA	420
Db	444 AGCTCTTCAAGAGAGACAGAGGTAGTCCGCTGCGCGAGTCAAGAGATTCTCA	503
Qy	421 AATTCTCAGCAGGGTCACTGAGGAGACAGGCTCTCCCAACCAAGCGGCGAGGAGT	480
Db	504 AATTCTCAGCAGG-----GGCTCTCCCAACCAAGCGGCGAGGAGGAGT	548
Qy	481 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	540
Db	549 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	608
Qy	541 CAGCAGTCAAG	600
Db	609 CAGCAGTCAAG	668
Qy	601 CGGGTTGAG	660
Db	669 CGGGTTGAG	728
Qy	661 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	720
Db	729 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	788
Qy	721 TCAAGGAG	780
Db	789 TCAAGGAG	848
Qy	781 GCCTTCTGGGCGGAGCTACCTGAGTGGCGCTCTGTCAGGAGCTTGGGGGCTGTCTG	840

Db	849 GCCTTCTGGGCGGAGCTACCTGAGTGGCGCTCTGCTGACAGGCTTGGCGGCGCTTCTCTG	908
Qy	841 ACTAGAGGCTTGGCGGAGAGAGCTGTGGGCGGAGAGGCTTGGCGGCTGCTGAGTGGAT	900
Db	909 ACTAGAGGCTTGGCGGAGAGAGCTGTGGGCGGAGAGGCTTGGCGGCTGCTGAGTGGAT	968
Qy	901 GAGGCTGAC 909	
Db	969 GAGGCTGAC 977	
RESULT 11		
LOCUS	BC013372	
DEFINITION	Homo sapiens death effector domain containing 2, mRNA (cDNA clone MGC:16414 IMAGE:3941907), complete cds.	
ACCESSION	BC013372	
VERSION	BC013372.2 GI:33872465	
KEYWORDS	MGC.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
AUTHORS	1 (bases 1 to 2005) Straubeberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toohily, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Muliahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.U., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettleman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skaleka, U., Smalins, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.	
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
PubMed	12477932	
REFERENCE	2 (bases 1 to 2005) Straubeberg, R.	
AUTHORS	Direct Submission	
JOURNAL	Submitted (31-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov On Aug 19, 2003 this sequence version replaced gi:15426522.	
COMMENT	Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Institute for Systems Biology http://www.systemsbio.org contact: amandareystemsbio.org Madan, Stephanie Rodriguez, Erin Helton, Mark Kettleman, Anuradha	
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAL, Plate: 25 Row: f Column: 8. Location/Qualifiers 1..2005 /organism="Homo sapiens" /mol_type="mRNA"	
SOURCE		

/db_xref="taxon:9606"
/clone="MGC:16414 IMAGE:3941907"
/cissue_type="lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: POTB7"
1. 2005
/gene="DDEDD2"
/note="synonym: FLAME-3"
/db_xref="LocusID:162989"
75_632
/gene="DDEDD2"
/codon_start=1
/product="DDEDD2 protein"
/protein_id="AAH13372.2"
/db_xref="GI:33872466"
/db_xref="LocusID:162989"
/translation="MALSGSTPAPCWEDECDLYGMLSHRMPEVVGQUTCELEL
LAFLEBAPGAGGLAPARSGLELLEBERGQCDENLKLGLQLKVLKRDHLPPL
ARRRRPVSPEYSGTSSSKRTGSCRRRRSSSSANSQSGSPPTKQRNRGRPS
GCARRRRRGAAPAPQ00SEPAQTFLL"

ORIGIN

Query Match 93.0%; Score 845.4; DB 9; Length 2005;
Best Local Similarity 98.0%; Pred. No. 1.2e-124;
Matches 893; Conservative 0; Mismatches 1; Indels 17; Gaps 3;

QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGGCTGGAGAGATGAGTGCCTGACTAC 60
DB 75 ATGGCGCTATCCGGGTGACCCCGCCGCTGGCTGGAGAGATGAGTGCCTGACTAC 134
QY 61 TACGGAGTCTGCTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
DB 135 TACGGAGTCTGCTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 194
QY 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCCCGCCGAGGCTTATGCC 180
DB 195 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCCCGCCGAGGCTTATGCC 254
QY 181 CGGGCCCGAGCGGCTTGAAGCTCTGCTGAGAGTGAAGCGCGGAGTGTGGCGGAG 240
DB 255 CGGGCCCGAGCGGCTTGAAGCTCTGCTGAGAGTGAAGCGCGGAGTGTGGCGGAG 314
QY 241 AGCAACTGCGGCTGCTGGGCAACTCTCTGCGCGTGTGGCCCGCCGACGACTGCTCCG 300
DB 315 AGCAACTGCGGCTGCTGGGCAACTCTCTGCGCGTGTGGCCCGCCGACGACTGCTCCG 374
QY 301 CACCTGGCGGCAAGCGGCGCGGCTGTCTTCAAGAGCTTATGCTATGCACTTCC 360
DB 375 CACCTGGCGGCAAGCGGCGCGGCTGTCTTCAAGAGCTTATGCTATGCACTTCC 434
QY 361 AGCTCTTCAAGAGAGAGAGAGGAGTGTGCGGCTGCGGCGGAGTGAAGAGTGTGCA 420
DB 435 AGCTCTTCAAGAGAGAGAGGAGTGTGCGGCTGCGGCGGAGTGAAGAGTGTGCA 494
QY 421 AATTTCAGAGAGGTCAGTGGAGACAGGCTCCCTCCCAACCAAGCGGACGCGGAGT 480
DB 495 AATTTCAGAGAGGTCAGTGGAGAGAGGCTCCCTCCCAACCAAGCGGACGCGGAGT 539
QY 481 CGGGCCCGGCGGCTGCTGCTGCAAGCGGCGGAGAGAGAGGAGGCGGCGGAGAG 540
DB 540 CGGGCCCGGCGGCTGCTGCTGCAAGCGGCGGAGAGAGAGGAGGCGGCGGAGAG 599
QY 541 CAGCAGTCAGAGAGCGGCGGCTGCTGCAAGAGAGAGTGAAGTGTGATCCGAGT 599
DB 600 CAGCAGTCAGAGAGCGGCGGAGAGTCTTCTTGAAGAGAGAGTGAAGTGTGATCCGAGT 659
QY 600 CCGGGTTCAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGTCCG 659
DB 660 CCGGGTTCAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGTCCG 719
QY 660 CGGGCCCGAGGCGGCTGCGGCGGAGTGTGAGAGTGTGAGAGAGTGTGAGTCCG 718

DB 720 GCGGCCCCAGAGCGCTGCGCGGAGAGTGTGTTGGGAGAGGACCGAGTGTCTC 779
QY 719 GCTCAAGGAGAGCTGAGGCTGTGTGTTGTGACATCAAGTTTCAAGCTTCTTATCTGG 778
DB 780 GCTCAAGGAGAGCTGAGGCTGTGTGTTGTGACATCAAGTTTCAAGCTTCTTATCTGG 839
QY 779 AGCCTTTCTGGGGGAGCTATCTGATGAGGAGGCTTGTGAGAGGCTTGGGGGAGTCC 838
DB 840 AGCCTTTCTGGGGGAGCTATCTGATGAGGAGGCTTGTGAGAGGCTTGGGGGAGTCC 899
QY 839 TGACTAGAGGCGCTGCGAGAGGAGTGTGAGGCGGAGGAGTGTGAGTGTGAGTGTG 898
DB 900 TGACTAGAGGCGCTTGCAGAGAGGCTGTGGCCGAGAGGCTGTTCCTGTGAGTGTG 959
QY 899 ATGAGGCTGAC 909
DB 960 ATGAGGCTGAC 970

RESULT 12
AX364851 1084 bp DNA linear PAT 15-FEB-2002
LOCUS AX364851
DEFINITION Sequence 2 from Patent WO0206315.
ACCESSION AX364851
VERSION AX364851.1 GI:18696740
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Mintz, L., Freilich, S. and Bernsteln, J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 2 24-JAN-2002;
CompuGen Ltd. (IL)

FEATURES
source location/Qualifiers
1..1084
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 90.8%; Score 825.8; DB 6; Length 1084;
Best Local Similarity 98.0%; Pred. No. 1.8e-121;
Matches 891; Conservative 7; Mismatches 4; Indels 7; Gaps 6;

QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGGCTGGAGAGATGAGTGCCTGACTAC 60
DB 84 ATGGCGCTATCCGGGTGACCCCGCCGCTGGCTGGAGAGATGAGTGCCTGACTAC 143
QY 61 TACGGAGTCTGCTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 120
DB 144 TACGGAGTCTGCTGCTTCAACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGC 203
QY 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCCCGCCGAGGCTTATGCC 180
DB 204 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCCCGCCGAGGCTTATGCC 262
QY 181 CGGGCCCGGAGCGGCTTGAAGCTCTGCTGAGAGTGAAGCGCGGAGTGTGGCGGAG 240
DB 263 CGGGCCCGGAGCGGCTTGAAGCTCTGCTGAGAGTGAAGCGCGGAGTGTGGCGGAG 322
QY 241 AGCAACTGCGGCTGCTGGGCAACTCTTGGCGGTGTGGCCCGCCGACGACTGCTCCG 300
DB 323 AGCAACTGCGGCTGCTGGGCAACTCTTGGCGGTGTGGCCCGCCGACGACTGCTCCG 382
QY 301 CACCTGGCGGCAAGCGGCGCGGCTGTCTTCAAGAGAGTGTGATGCTATGCACTTCC 360
DB 383 CACCTGGCGGCAAGCGGCGCGGCTGTCTTCAAGAGAGTGTGATGCTATGCACTTCC 440
QY 361 AGCTCTTCAAGAGAGAGAGGAGTGTGCGGCTGCGGCGGAGTGAAGAGTGTGCA 420
DB 441 AGCTCTTCAAGAGAGAGAGGAGTGTGCGGCTGCGGCGGAGTGAAGAGTGTGCA 500

[illegible]

RESULT 13				
LOCUS	AX364852	1106 bp	DNA	linear
DEFINITION	Sequence 3 from Patent WO0206315.			
ACCESSION	AX364852			
VERSION	AX364852.1	GI:18696741		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 Mintz, L., Freilich S. and Bernstein, J.			
AUTHORS	Novel nucleic acid and amino acid sequences			
TITLE	Patent: WO 0206315-A 3 24-VAN-2002;			
JOURNAL				

FEATURES	Location/Qualifiers
source	1..1106
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"

ORIGIN	
Query Match:	90.8%
Score	825.8
DB	6
Length	1106
Best Local Similarity	98.0%
Pred. No.	1.8e-121
Matches	891
Conservative	7
Indels	4
Gaps	7
	6

Oy 1 TTGGGGCATCCGGGTGCACCCGGGCCCCGGTGGGGAGAGAGATGCTGGACAC 60
 Db 106 ATGGGCGTATCCGGGTGCACCCGGGCCCCGGTGGGGAGAGATGCTGGACAC 165
 Oy 61 TACGGAGTCTGTGGCTTACCGGATATGTTCCAGAGTGTTGGGGCGGCAACTGACCGAGTGC 120
 Db 166 TACGGAGTCTGTGGCTTACCGGATATGACAGAGTGTTGGGGCGGCAACTGACCGAGTGC 225

OY	121	AAGCTGAGAGCTCCCTGCGCCCTTCTGTGTGATAGAGCTCCTGGGCGCCCGGAGGCTTATACC	180
Db	226	TAAGCTGAGAGCTCCCTGCGCCCTTCTGTGTGATAGAGCTCCTGGGCGCCCGGAGG--TTATACC	284
OY	181	CGGGCCCGCAACGGGCTTAGAGCTCTGCTGTGAGGTGAGCGCCCGCGGACAGTGCAG	240
Db	285	CGGGCCCGCAACGGGCTTAGAGCTCTGCTGTGAGGTGAGCGCCCGGACAGTGCAG	344
OY	241	AGCAACTGTGGGCTGTGGGGCAACTCTGTGGCTGTGGCCCGCACGACTGTGCTCG	300
Db	345	AGCAACTGTGGGCTGTGGGGCAACTCTGTGGCTGTGGCCCGCACGACTGTGCTCG	404
OY	301	CACCGGCGGCAAGCGGCGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGGACCTCC	360
Db	405	CACCTGGCGGCA--AGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGGACCTCC	462
OY	361	AGCTTCTCAAGAGAGACAGAGGGTAGTCCGTGCGCGTGGGACAGTCAAGCATTTGCA	420
Db	463	AGCTTCTCAAGAGAGACAGAGGGTAGTCCGTGCGCGTGGGACAGTCAAGCATTTGCA	522
OY	421	AATTCTCAGCAGGGGTCAGTGGGAGACAGGCTCCCCCCTAACCAAGCGGACGGCGAGT	480
Db	523	AATTCTCAGCAGGGGTCAGTGGGAGACAGGCTCCCCCCTAACCAAGCGGACGGCGAGT	581
OY	481	CGGGGCGCGCCAGATGGTGTGTCAGAGCGGCGGAGAGGGGCCCCAGCGCACCCAG	540
Db	582	CGGGGCGCGCCAGATGGTGTGTCAGAGCGGCGGAGAGGGGCCCCAGCGCACCCAG	640
OY	541	CAGCAGTCAGAGCCCGCGCACACTTCTCTGAAGGCAAAAGTACCTGTGACATCCGCTC	600
Db	641	CAGCAGTCAGAGCCCGCGCACACTTCTCTGAAGGCAAAAGTACCTGTGACATCCGCTC	700
OY	601	CGGGTTGAGAGAGTACTGTGACATAGGGGCAAGCTTGTGAGCAAGGGGTGTGACATCCGG	660
Db	701	CGGGTTGAGAGAGTACTGTGACATAGGGGCAAGCTTGTGAGCAAGGGGTGTGACATCCGG	760
OY	661	CGGCCCCAGGCGCTGAGCGCGGCAAGCTGTGTTTGGGCAAGGCCACCGCAGTGTGCGC	720
Db	761	CGG--CCGAGGCGCTGTGGCGGCAACKGGACGTGTTTGGGCAAGGCCACCGCAGTGTGCGC	819
OY	721	TCAAGGGACCTGGGCTCTGTGTGTTTGTGACATCAAGTCTCAGAGCTTCTATCTGGAC	780
Db	820	TCAAGGGACCTGGGCTCTGTGTGTTTGTGACATCAAGTCTCAGAGCTTCTATCTGGAC	879
OY	781	GCTTCTGGGGGAGACTACTGAGTGGGGCCCTGTGCGAGGCGCTCGGGGGGCGTGTCTCG	840
Db	880	GCTTCTGGGGGAGACTACTGAGTGGGGCCCTGTGCGAGGCGCTCGGGGGGCGTGTCTCG	939
OY	841	ACTAGGCGCTTGCAGAGAGCTGTGTGGGCGGGAGGCTGTTGCGCTGTGCTGAGTGTGAT	900
Db	940	ACTAGGCGCTTGCAGAGAGCTGTGTGGGCGGGAG--GCTGTTGCGCTGTGCTGAGTGTGAT	998
OY	901	GAGGCTGAC 909	
Db	999	GAGGCTGAC 1007	

RESULT	14
AF457576	
LOCUS	
DEFINITION	993 bp mRNA linear ROD 10-APR-2002
ACCESSION	AF457576
VERSION	AF457576.1 GI:20126795
	(Flame-3) mRNA, complete cds.

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
Mus musculus (house mouse)	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; 1 (bases 1 to 993)	Zhan, Y., Hegde, R., Srinivasula, S.M., Fernandes-Alnemri, T. and Alnemri, E.S.	Death effector domain-containing proteins DEDD and FLAME-3 form

ORIGIN

Query Match 77.1%; Score 701; DB 10; Length 993;
Best Local Similarity 86.8%; Pred. No. 1,1e-101;
Matches 799; Conservative 0; Mismatches 110; Indels 12; Gaps 2;

```
QY      1 ATGGCGCTATCCGGGTCGACCCCGCCCGCTGTCTGGAGAGATGATGCTGCTGACTAC 60
        |||
DB       1 ATGGCGCTATCCGGGTCGACCCCGCCCGCTGTCTGGAGAGATGATGCTGCTGACTAC 60
QY      61 TAGCGGATGCTGTGCTTCAACCTATGTTCCAGAGTGTGGCGGGCAACTGACCGAGTGC 120
        |||
DB       61 TAGCGGATGCTGTGCTTCAACCTATGTTCCAGAGTGTGGCGGGCAACTGACCGAGTGC 120
QY      121 GAGCTGAGCTCTGAGCTTCTCTGTGATGAGGCTCTTGCGCCCGCGAGGCTTAGCC 180
        |||
DB       121 GAGCTGAGCTCTGAGCTTCTCTGTGATGAGGCTCTTGCGCCCGCGAGGCTTAGCC 180
QY      181 CGGCGCCGCGAGCGGCTTAGAGCTCTGTGAGCTGAGCGCGCGGGCAGTGCAG 240
        |||
DB       181 CGGCGCCGCGAGCGGCTTAGAGCTCTGTGAGCTGAGCGCGCGGGCAGTGCAG 240
QY      241 AGCAACCTGCGGCTGTGAGGCAACTCTGTGCGGCTGTGCGCCGACGACTCTGTCCA 300
        |||
DB       241 AGCAACCTGCGGCTGTGAGGCAACTCTGTGCGGCTGTGCGCCGACGACTCTGTCCA 300
QY      301 CACCTGGCGCGAGCGGCGGCGGCGAGTGTCTCCAGAAAGCTATAGCTATGGCACTCC 360
        |||
DB       301 CACCTGGCGCGAGCGGCGGCGGCGAGTGTCTCCAGAAAGCTATAGCTATGGCACTCC 360
QY      361 AGC---TCTTCAAGAGAGACAGAGGTAAGTCCGTGCGCGTCTGAGTCAAGCAGTTCT 417
        |||
DB       361 AGCTCTTCTTCCAGAGAGACAGAGGTAAGTCCGTGCGCGTCTGAGTCAAGCAGTTCT 420
QY      418 GCAAAATTCAGAGAGGTCATGTGGAGACAGGCTTCCCCCAACCAAGCGGAGGCGG 477
        |||
DB       421 TCGGACTCTCTCAGAGTCAAGTGGAGACAGGCTTCCCCCAACCAAGCGGAGGCGG 480
QY      478 AGTCGGGGCGGCGCAGTGTGTGTCAGAGCGGCGGAGAGAGGGGCGCCAGCGCGC--- 533
        |||
DB       481 AGTCGGGGCGGCGCCTTAGAGTGTGTCGCGGAGCGGCGGAGAGAGGCGGCTTTCAGGCTCG 540
QY      534 -----ACCCGAGCAGCAGTCAAGCCCGCAGACTTCTCTGAAAGGCMAAGTGAACCTGT 588
        |||
DB       541 CAGCAGACACAGCAGACACAGAGTGTGGGCAAGGCCCTCTCAGAAAGCAAGTGAACCTGT 600
QY      589 GACATCCGGCTCCGGGTTCCAGCAGAGTACTGCGAGCATGGGCCAGCTTGGAGCAGGGC 648
        |||
DB       601 GACATCCGGCTCCGGGTTCCGGGTCAGAGTACTGCGAGCATGGGCCAGCTTGGAGCAGGGT 660
QY      649 GTGGCATCCCGGCGGCGCCAGAGCGCTGGCGGCGGAGCTGAGCGTGTGGGCGAGGCCACC 708
        |||
DB       661 GTGGCGTCCCGGAGACCTCAGAGCGTGGCGGCGGAGCTGAGTGTGTGGACAAAGCCACT 720
QY      709 GCAGTGTGCGCTCAAGGACCTGAGCTCTGTGTTTGTGACATCAAGTTCTCAGAGCTC 768
        |||
DB       721 GCTGTGTGCGGCTCAAGGACCTGAGGCTCTGTGTGTGTGACATCAAGTTCTCAGAGCTG 780
QY      769 TCCATCTGAGCGCTTCTGGGGGAGACTACTGAGTGGCGCTGTGCGAGGCCCTGTGCG 828
        |||
DB       781 TCCATCTGAGCGCTTCTGTGGGGGAGACTACTGAGTGGCGCTGTGCGAGGCCCTTTCG 840
QY      829 GCGGTGTTCTGAGTGAAGGCGCTGAGAGGCTGTGGGCGGAGGCGTGTGCGCTGCTG 888
        |||
DB       841 GCGGTGTTCTGAGTGAAGGCTTTCAGGAGGCGCTGTGGGCGGAGGCGTGTGCGCTGCTG 900
QY      889 GTCAGTGTGATGAGGCTGAC 909
        |||
DB       901 GTCAGTGTGAGCAGAGGCTGAC 921
```

Search completed: March 23, 2005, 13:22:32
Job time : 4099.8 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 01:03:41 ; Search time 26.0458 Seconds
(without alignments)
5210.505 Million cell updates/sec

Title: US-10-030-271-1
Perfect score: 1716
Sequence: 1 atgagcgtaccgcggtcgcac.....tcagtgatgagtcgcac 909

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=x1p
-Q=/cgn2_1/USFNO_spool_p/US10030271/runat_21032005_153459_26492/app_query.fasta_1.3150
-DB=Issued_Patents_AA -QFWT=faetan -SUFFIX=rai -MINMATCH=0.1 -LOOPCTL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=PCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10030271 @CGN 1.1 81 @runat_21032005_153459_26492 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614.5	35.8	318	3	US-08-859-167-4
2	614.5	35.8	318	3	US-09-109-273-4
3	614.5	35.8	318	3	US-09-276-993-4
4	614.5	35.8	318	4	US-09-723-450-4
5	614.5	35.8	318	4	US-09-733-167A-1
6	608.5	35.5	318	3	US-08-859-167-6
7	608.5	35.5	318	3	US-09-109-273-6
8	608.5	35.5	318	3	US-09-276-993-6
9	608.5	35.5	318	4	US-09-723-450-6
10	608.5	35.5	318	4	US-09-733-167A-3
11	422	24.6	210	4	US-09-733-167A-6
12	225.5	13.1	114	4	US-09-733-167A-5

C 13	221.5	13.5	369	4	US-09-252-991A-25394	Sequence 25394, A
C 14	217	13.2	1034	4	US-09-252-991A-28921	Sequence 28921, A
C 15	210.5	12.9	309	4	US-09-252-991A-25386	Sequence 25386, A
C 16	210	12.8	273	3	US-09-252-991A-23651	Sequence 23651, A
C 17	210	12.8	504	3	US-09-219-849-3	Sequence 3, Appl1
C 18	210	12.8	561	1	US-08-642-255-52	Sequence 52, Appl1
C 19	209	12.8	1706	4	US-09-252-991A-31760	Sequence 31760, A
C 20	205	12.5	1411	4	US-09-252-991A-28408	Sequence 28408, A
C 21	204	11.9	501	4	US-09-252-991A-22799	Sequence 22799, A
C 22	204	11.9	630	4	US-09-252-991A-25626	Sequence 25626, A
C 23	204	12.5	869	4	US-09-252-991A-22290	Sequence 22290, A
C 24	201.5	11.7	957	4	US-09-252-991A-20408	Sequence 20408, A
C 25	201	11.7	456	4	US-09-252-991A-19417	Sequence 19417, A
C 26	201	12.3	720	3	US-09-219-849-4	Sequence 4, Appl1
C 27	201	12.3	777	1	US-08-642-255-53	Sequence 53, Appl1
C 28	200	12.2	315	4	US-09-252-991A-20553	Sequence 20553, A
C 29	199	11.6	476	4	US-09-252-991A-27078	Sequence 27078, A
C 30	198	12.1	1418	4	US-09-252-991A-32367	Sequence 32367, A
C 31	197.5	12.1	491	4	US-09-489-039A-7836	Sequence 7836, Ap
C 32	197	11.5	1057	3	US-08-931-820-4	Sequence 4, Appl1
C 33	196	12.0	973	4	US-09-252-991A-21386	Sequence 21386, A
C 34	195	11.9	957	4	US-09-252-991A-20408	Sequence 20408, A
C 35	194.5	11.3	345	4	US-09-252-991A-18076	Sequence 18076, A
C 36	194	11.8	477	4	US-09-252-991A-19631	Sequence 19631, A
C 37	194	11.3	1078	3	US-08-963-825-21	Sequence 21, Appl1
C 38	194	11.3	1078	3	US-09-500-811-21	Sequence 21, Appl1
C 39	194	11.3	1078	3	US-09-570-573-21	Sequence 21, Appl1
C 40	194	11.3	1078	3	US-09-548-608-21	Sequence 21, Appl1
C 41	193.5	11.3	377	4	US-09-252-991A-31793	Sequence 31793, A
C 42	193.5	11.3	398	4	US-09-252-991A-21102	Sequence 21102, A
C 43	193	11.8	234	1	US-08-642-255-51	Sequence 51, Appl1
C 44	192.5	11.8	472	4	US-09-252-991A-31978	Sequence 31978, A
C 45	192.5	11.8	550	4	US-09-252-991A-21295	Sequence 21295, A

ALIGNMENTS

RESULT 1
US-08-859-167-4
Sequence 4, Application US/08859167
Patent No. 6037461
GENERAL INFORMATION:
APPLICANT: Alnemir, Enad S.
TITLE OF INVENTION: RADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,167
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 318 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-859-167-4

```

Alignment Scores:	
Pred. No.:	1,73e-37
Score:	614.50
Percent Similarity:	62.79%
Best Local Similarity:	44.85%
Query Match:	35.81%
DB:	3
Length:	318
Matches:	139
Conservative:	54
Mismatches:	91
Indels:	21
Gaps:	6

US-10-030-271-1 (1-909) X US-08-859-167-4 (1-318)

[illegible]

Db 280 SerLeuLygGlnIaValaGlyHisGlnIatIleuValaAsnValaAspGluGlu 299
Qy 907 GAC 909
Db 300 Asp 300

```

1  RESULT 2
2  US-09-109-273-4
3  : Sequence 4, Application US/09109273
4  : Patent No. 6063760
5  : GENERAL INFORMATION:
6  : APPLICANT: Alnemri, Emad S.
7  : APPLICANT: Fernandez-Alnemri, Teresa
8  : TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
9  : TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
10 : TITLE OF INVENTION: OF MAKING THE SAME
11 : NUMBER OF SEQUENCES: 17
12 : CORRESPONDENCE ADDRESS:
13 : ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760-ris
14 : STREET: One Liberty Place, 46th floor
15 : CITY: Philadelphia
16 : STATE: PA
17 : COUNTRY: USA
18 : ZIP: 19103
19 : COMPUTER READABLE FORM:
20 : MEDIUM TYPE: Floppy disk
21 : COMPUTER: IBM PC compatible
22 : OPERATING SYSTEM: WINDOWS
23 : SOFTWARE: Wordperfect
24 : CURRENT APPLICATION DATA:
25 : APPLICATION NUMBER: US/09/109,273
26 : FILING DATE:
27 : CLASSIFICATION:
28 : PRIOR APPLICATION DATA:
29 : APPLICATION NUMBER: 08/859,167
30 : FILING DATE:
31 : ATTORNEY/AGENT INFORMATION:
32 : NAME: Deluca, Mark
33 : REGISTRATION NUMBER: 33,229
34 : REFERENCE/DOCKET NUMBER: TUI-
35 : TELECOMMUNICATION INFORMATION:
36 : TELEPHONE: (215) 568-3100
37 : TELEFAX: (215) 568-3439
38 : INFORMATION FOR SEQ ID NO: 4:
39 : SEQUENCE CHARACTERISTICS:
40 : LENGTH: 318 amino acids
41 : TYPE: amino acid
42 : TOPOLOGY: linear
43 : MOLECULE TYPE: protein
44 : US-09-109-273-4

```

Alignment Scores:	
Pred. No.:	1,73e-37
Score:	614.150
Percent Similarity:	62.79%
Best Local Similarity:	44.85%
Query Match:	35.81%
DB:	3
Gaps:	6
Length:	318
Matches:	64
Conservative:	54
Mismatches:	91
Indels:	21
Gaps:	6

US-10-030-271-1 (1-909) X US-09-109-273-4 (1-318)

QY 34 TGGAGAGAGATGATGCTGGACTGACCTACTACCGGATGCTGTGGCTTACCGGATTTTGAG 93
 |||||
 Db 12 TTPRGGLGGLuhhlsclgyluglnsluhhlsclgyleutyrsleluhHsARMeTPhleAsp 31
 QY 94 GTGTGTGGGGGGGAATGACCGAGTGGACGCTCGTGGCCTTTCTG---CTTGAT 150
 |||||
 Db 32 lIeValGlyThrhIsleuthrhIsaRgAspValaRgValaLeuSerPheLeuPheValaAsp 51
 QY 151 GAAGGCTCTGGGCGCGCGGAGGACTTAAAGCCGGGCGCGGAGCGGACTTAAGACTCTCTG 210
 |||||
 Db 52 ValIleAspAspPhIsclnARgclleu-----lIeRgAsnGlyARgAspPheLeu 69

```

QY 211 GAGTGGAGCCGCGGCGAGTGGCGGAGCAACTGGCGCTGTGGGCACTCTG 270
    |||||
Db 70 AlaleugluArgInglYArgCyAspIuSerAspIuSerAspIuValIleuInleu 89
QY 271 CGCGTGTGGCCCGGCGAGCACTGTCTGGCAGCTGGGCGGAGCGGCGGCGAGT 330
    |||||
Db 90 ArgIleIleThrArgHisAspIuSerProTyValIleuIleuValArgAlaVal 109
QY 331 TCTCCA-----GAAGCTATAGTATGGACCTCC-----AGCTCTCAAG 372
    |||
Db 110 CysProAspIuValAspIuSerTyIleuIngluInuThrSerIleArgTyValIleuProArg 129
QY 373 AGGACAGAGGTAAGTGGCGGCGGCGAGTGGCGAGTCTGCAATTCTCAGAG 432
    |||
Db 130 AlaleuSerAspProgluProArgProProgluInProSerIleuValProProIleuTy 149
QY 433 GGTGAG-----TGGAGACAGAGCTCCCGGCGAGCAAGCGGAGCGGAGTGGGCG 486
    |||
Db 150 ProValValCysCysProThrSerIleProgluMetCysSerIleuArgProAlaArgIle 169
QY 487 CGGCGCAAGTGTGGTGGCGGAGCGGCGGAGGCGGCGGCGGCGGCGGCGGAG 546
    |||
Db 170 ArgAlaThrIleuGlySerGlnArgIuArgArg----- 180
QY 547 TCAGAGCCGCGGCGAGCTTCTCTGAGAGGCAAGTGAAGTCTGTGACATCCGAGTCCGGGT 606
    |||
Db 181 ---LysSerValThrProAspProIuSerIuSerIuThrCysAspIleArgIuSerVal 199
QY 607 CGAGCAGAGTACTGGAGCAATGGCGGCGAGCTTGGAGCGAGGCGGTGGCATCCCGGCGGCG 666
    |||
Db 200 ArgAlaGluTyArgCysGlnHisGluIuThrAlaIleuIngluAspValPheSerAsnIuSerGln 219
QY 667 CAGGCGCTGGGCGGCGAGCTGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCAGG 726
    |||
Db 220 AspProIuSerGluArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleuIuSerArg 239
QY 727 GACCTGGGCTGTGGTGTGGACATCAAGTCTCAGAGCTCTCTATCTGGAGCGCTTC 786
    |||||
Db 240 AspIuGlySerIleIleCysAspIleIleuPheSerIleuIleuThrTyIleuAspAlaPhe 259
QY 787 TGGGCGCACTACTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTCTGACTAG 846
    |||
Db 260 TrpArgAspTyIleAsnGlySerIleuGlnAlaIleuValIleuPheIleuThrAsp 279
QY 847 GCCCTGGAGAGGCGTGGGCGGCGGAGCTGTGCCCTGGCTGTGAGTGTGATAGCT 906
    |||
Db 280 SerIuIuSerGlnAlaValGlyHisGlnAlaIleuIleuValAsnValAspIuGlu 299
QY 907 GAC 909
    |||
Db 300 Asp 300

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,993
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
FAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-276-993-4

Alignment Scores:
Pred. No.: 1,73e-37 Length: 318
Score: 614.50 Matches: 135
Percent Similarity: 62.79% Conservative: 54
Best Local Similarity: 44.85% Mismatch: 51
Query Match: 35.81% Gaps: 21
DB: 3

US-10-030-271-1 (1-909) x US-09-276-993-4 (1-318)
QY 34 TGGAGAGAGATGAGTGCCTGAGTAAAGGAGTGTGCTTACCGTATGTTGAG 93
    |||
Db 12 TrpProgluIngluHisGluIngluInuIleuIleuTyIleuIleuIleuIleuIleu 31
QY 94 GTGGTGGGCGGCGGCAAGTGAAGAGTGGAGTGGAGTCTGGGCGCTTCTG---CTGAT 150
    |||
Db 32 IleValGlyIleuHisIleuThrHisIleuArgAspValArgValIleuSerPheIleuPheValAsp 51
QY 151 GAGGCTCTGGGCGGCGGCGGAGCTTAAAGCCGCGGCGGCGGCGGCGGCGGCTAGCTCTG 210
    |||
Db 52 ValIleAspAspHisGluArgIleu-----IleArgAsnGlyArgAspPheIleu 69
QY 211 GAGTGGAGCGCGGCGGCGAGTGGCGGAGCAACTGGCGCTGTGGGCGCACTCTG 270
    |||||
Db 70 AlaleugluArgInglYArgCyAspIuSerAspIuSerAspIuValIleuInleu 89
QY 271 CGCGTGTGGCCCGGCGAGCACTGTCTGGCAGCTGGGCGGAGCGGCGGCGGCGGAGT 330
    |||||
Db 90 ArgIleIleThrArgHisAspIuSerProTyValIleuIleuValArgAlaVal 109
QY 331 TCTCCA-----GAAGCTATAGTATGGACCTCC-----AGCTCTCAAG 372
    |||
Db 110 CysProAspIuValAspIuSerTyIleuIngluInuThrSerIleArgTyValIleuProArg 129
QY 373 AGGACAGAGGTAAGTGGCGGCGGCGAGTGGCGAGTCTGCAATTCTCAGAG 432
    |||
Db 130 AlaleuSerAspProgluProArgProProgluInProSerIleuValProProIleuTy 149
QY 433 GGTGAG-----TGGAGACAGAGCTCCCGGCGAGCAAGCGGAGCGGAGTGGGCG 486
    |||
Db 150 ProValValCysCysProThrSerIleProgluMetCysSerIleuArgProAlaArgIle 169
QY 487 CGGCGCAAGTGTGGTGGCGGAGCGGCGGAGGCGGCGGCGGCGGCGGCGGAGTGGGCG 546
    |||
Db 170 ArgAlaThrIleuGlySerGlnArgIuArgArg----- 180
QY 547 TCAGAGCCGCGGCGAGCTTCTCTGAGAGGCAAGTGAAGTCTGTGACATCCGAGTCCGGGT 606
    |||

```


ORGANISM: Homo sapiens
US-09-733-167A-1

Alignment Scores:

Score:	1.73e-37	Length:	318
Percent Similarity:	614.50	Matches:	135
Best Local Similarity:	62.79%	Mismatches:	54
Query Match:	44.85%	Indels:	21
DB:	4	Gaps:	6

US-10-030-271-1 (1-909) x US-09-733-167A-1 (1-318)

```

QY 34 TGGAGAGAGATGATGCTGCTGACTACTACCGAGATGCTGCTGCTTCAACCGTATGTTGAG 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 TtrProgluInHtIsGlyGluInGlnHtIsGlyLeuTyrsrIleuHtIsArgMetPheAsp 31
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 GTGTGGGCGGCGCACTGACCGAGTGCAGCTGCAGCTCTGCGCTTTCTG--CTGGAT 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 IleValGlyThrHtIsLeuThrHtIsArgAspValArgValIleuSerPheLeuPheValAsp 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 GAGCTCTGCGCGCGCGGAGGCTAGCGCGGCGCGGAGCGGCTAGAGCTCTGCTGCT 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 52 ValIleAspAspHtIsGlyArgGlyLeu-----IleArgAsnGlyArgAspPheLeuLeu 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 GAGCTGAGCGCGCGCGGAGTGCAGCGGAGCAACCTGCGGCTGCTGGGCAACTCTG 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 AlaleuGluArgGlnGlyArgCyAspAspIleuSerAsnPheArgGlnValIleuGlnLeuLeu 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CGCGTCTGCGCGCGCGCACTGCTGCGCACTGCGCGCGCAACGCGCGCGCGCACTG 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 ArgIleIleThrHtIsArgHtIsAspLeuLeuProTyValThrLeuTyArgArgValAlaVal 109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 TCTGCA-----GAAGCTATAGCTATGCACTCTCC-----AGCTCTTCAAG 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 CyAspProAspLeuValAspTySerTyrLeuGlnIleuThrSerIleArgTyValThrProArg 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 373 AGGACAGAGGAGTACTGCGCTGCGCGGAGTGCAGTGCAGTCTGCAATTTCTCAGCAG 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 AlaleuSerAspProGluProArgProProGlnProSerIleTyThrValProProHtIsTyr 149
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 433 GGTGAG-----TGGAGACAGGCTCCCGCCCAACCAAGCGCGAGCGGCGAGTGGGGC 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 ProValValCyAspProThrSerGlyProGlnMetCySerIleTyArgProAlaArgGly 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 487 CGGCGCGAGTGTGCTGCGCGAGCGCGCGGAGGCGCGCGCGCGCGCGCGCGAGCAG 546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 ArgAlaThrLeuGlySerGlnArgIleArgValArgArg----- 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 TCAGAGCGCGCGCAACCTTCTCTGAAGGCAAAAGTGAACCTGATCCGGCTCGGGGT 606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 ---lysSerValThrProAspProIleGlySerGlnThrCyAspIleArgIleuAlaVal 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 CGAGCAAGTACTGCGAGCATGGCGCACTTGAAGAGCGGCGTGCAGTCCCGCGCGCC 666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 ArgAlaGluTyrCyGlnHtIsGlyThrAlaLeuGlnGlyAsnValPheSerAsnIleGln 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 667 CAGGCGCTGCGCGCGGAGTGCAGTCTTGGGAGAGCGCACCGAGTCTGCGCTCAAG 726
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 AspProLeuGluArgGlnPheGluArgPheAsnGlnAlaAsnThrIleLeuIleSerArg 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 727 GAGCTGAGCTGTGATTTGATGACATCAAGTTTCAAGCTCTCTATCTGAGAGCGCTTC 786
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 AspLeuGlySerIleIleCyAspIleTyPheSerGlyLeuThrTyIleuAspAlaPhe 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 787 TGGGAGCACTACTGAGTGGCGCGCTGTGACAGCGCTTGGCGGCGGTTGCTTGAAG 846
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 TtrArgAspTyIleAsnGlySerIleuGlnAlaLeuTyGlyValPheIleThrAsp 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 847 GCGCTGAGAGAGCTGTGCGCGGAGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 SerIleuTyGlnAlaValGlyHtIsGlyAlaIleTyLeuLeuValAsnValAspGlu 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 907 GAC 909
Db 300 Asp 300

RESULT 6

US-08-859-167-6

; Sequence 6, Application US/08859167

; Patent No. 6037461

; GENERAL INFORMATION:

; APPLICANT: Alnemri, Emad S.

; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF

; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS

; TITLE OF INVENTION: OF MAKING THE SAME

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s

; STREET: One Liberty Place, 46th floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: WINDOWS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/859,167

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Deluca, Mark

; REGISTRATION NUMBER: 33,229

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3439

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 318 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-859-167-6

Alignment Scores:

Pred. No.:	4.79e-37	Length:	318
Score:	608.50	Matches:	131
Percent Similarity:	62.13%	Conservative:	56
Best Local Similarity:	43.52%	Mismatches:	93
Query Match:	35.46%	Indels:	21
DB:	3	Gaps:	5

US-10-030-271-1 (1-909) x US-08-859-167-6 (1-318)

```

QY 34 TGGAGAGAGATGATGCTGCTGACTACTACCGAGATGCTGCTGCTTCAACCGTATGTTGAG 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12 TtrProgluInHtIsGlyGluInGlnHtIsGlyLeuTyrsrIleuHtIsArgMetPheAsp 31
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 GTGTGGGCGGCGCACTGACCGAGTGCAGCTGCAGCTCTGCGCTTTCTG--CTGGAT 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 IleValGlyThrHtIsLeuThrHtIsArgAspValArgValIleuSerPheLeuPheValAsp 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 GAGCTCTGCGCGCGCGGAGGCTTAGACCGGCGCGGAGCGGCTTGAAGCTCTGCTG 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 52 ValIleAspAspHtIsGlyArgGlyLeu-----IleArgAsnGlyArgAspPheLeuLeu 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 GAGCTGAGCGCGCGCGGAGTGCAGCGGAGCAACCTGCGGCTGCTGGGCAACTCTG 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 AlaleuGluArgGlnGlyArgCyAspAspIleuSerAsnPheArgGlnValIleuGlnLeuLeu 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 CGCGTCTGCGCGCGCGCACTGCTGCGCACTGCGCGCGCAACGCGCGCGCGCGCACTG 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


QY 667 CAGGCGCTGGCGCCGACACTGCTGTTGGGCGAGCCACCGCAGTCTGCTCAAG 726
 Db 220 AAspProleuGlnArgGlnPheGlnPheAsnGlnAlaAsnThrLleuLysSerArg 239
 QY 727 GACCTGGGCTGCTGGTTTGACATCAAGTTCACAGCTCTCTATCTGAGCCCTTC 786
 Db 240 AAspLeuGlySerLleLleCysAspLleLysPheSerGluLeuThrTyLLeuAspAlaPhe 259
 QY 787 TGGGGCGACTACTGAGTGGCGCCCTGCTGACAGCCCTGGCGGCGCTTCCTCACTGAG 846
 Db 260 TrpArgAspTrpLleAsnGlySerLeuLeuGlnAlaLeuLysGlyValPheLleThrAsp 279
 QY 847 GCCCTGCGAGAGCTGTGGCGCCGAGGCTGTTGCGCTGCTGTCAGTGTGATGAGCT 906
 Db 280 SerLeuLysGlnAlaValGlyHisGlnLleLleLysLeuLeuValAsnValAspGluGlu 299
 QY 907 GAC 909
 Db 300 Asp 300

RESULT 8

US-09-276-993-6
 ; Sequence 6, Application US/09276993
 ; Patent No. 6207801

; GENERAL INFORMATION:
 ; APPLICANT: Alnemrl, Emad S.
 ; APPLICANT: Fernandez-Alnemrl, Teresa
 ; TITLE OF INVENTION: PADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
 ; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
 ; TITLE OF INVENTION: OF MAKING THE SAME
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Woodcock, Maeburn, Kurtz, Mackiewicz & No. 6207801r1s
 ; STREET: One Liberty Place, 46th floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: WINDOWS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/276,993
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/859,167
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Deluca, Mark
 ; REGISTRATION NUMBER: 33,229
 ; REFERENCE/DOCKET NUMBER: TJU-
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 318 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-276-993-6

Alignment Scores:

Pred. No.: 4.79e-37 Length: 318
 Score: 608.50 Matches: 131
 Percent Similarity: 62.13% Conservative: 56
 Best Local Similarity: 43.52% Mismatches: 93
 Query Match: 35.46% Indels: 21
 DB: 3 Gaps: 5

US-10-030-271-1 (1-909) x US-09-276-993-6 (1-318)

QY 34 TGGAGAGAGATGATGCTGCTGACTGATCAAGAGATGCTGTCTTACCGGATGTCAG 93
 Db 12 TrpProGlnGlnArgGlyGlnGlnGlnHisLysLeuTySerLleuHisArgMetPheAsp 31
 QY 94 GTGGTGGCGGGGCACTGACCGAGTGGAGCTGAGCTCTGCGCTTTCTG--CTGAT 150
 Db 32 LLeValGlyThrHisLleuThrHisArgAspValArgValLysSerPheLeuPheValAsp 51
 QY 151 GAGGCTCTGGCGCGCGCGAGGCTTAGCCCGGCGCGAGCGGCTTAGAGCTCTGTG 210
 Db 52 ValLLeAspAspHisGlnArgGlyLeu-----LLeArgAsnGlyArgAspPheLeuLeu 69
 QY 211 GAGCTGAGCGCGCGCGGAGTGGCGGAGCAAGCTGCGGCTGTGGGCAACTCTG 270
 Db 70 AlaLeuGlnArgGlnGlyArgCysAspGluSerAsnPheArgGlnValLeuGlnLeuLeu 89
 QY 271 CGCGTGTGCGCCCGCAAGACTGTGCGGCACTGCGCGCGCGCAAGCGCGCGCGACTG 330
 Db 90 ArgLleLleThrArgHisAspLeuLeuProTyValThrLeuLysArgAlaVal 109
 QY 331 TCTCCA-----GAACGTATAGCTATGCGACCTTCGACTTTCAAGAGACAGAG 381
 Db 110 CysProAspLeuValAspLysTyLLeuGlnGlnThrSerLLeArgTyValThrProArg 129
 QY 382 GTTAGTGTGCGCGTCCCGCGGAGTCAAGAGTTCGAAATTCACAGAGGTCAGTGG 441
 Db 130 AlaLeuSerAspProGluProArgProProGlnProSerLysThrValProProHisTy 149
 QY 442 GAGACAGGCTCCCGCCCAACCAAG-----CGCAGCGCGGAGTGTGCGGC 486
 Db 150 ProValValCysCysProThrSerGlySerGlnMetCysSerLysArgProAlaArgGly 169
 QY 487 CGGCGCCAGTGTGTGTCAGACGCGCGGCGAGAGGCGCGCCAGCGCGACCCAGCAGAG 546
 Db 170 ArgThrThrLeuGlySerGlnArgLysArgArg----- 180
 QY 547 TCAGAGCCCGCGCAACCTCTCTGAAGGCGAAAGTGAATGATCGGCTCGGCTT 606
 Db 181 ---LysSerValThrProAspProLysGlnLysGlnThrCysAspLLeArgLeuArgVal 199
 QY 607 CGAGCAGATGACTGCGAGCATGCGCGCAGCTTGAGAGAGGCGGTGCGATCCCGCGCGCC 666
 Db 200 ArgAlaGlnTyTyCysGlnHisGlnThrAlaLeuGlnGlnGlnValAsnValPheSerAsnLysGln 219
 QY 667 CAGGCGCTGGCGCGGACGCTGCTTGGGCGACCGCAGCGAGTGTGCTCAAG 726
 Db 220 AAspProleuGlnArgGlnPheGlnPheAsnGlnAlaAsnThrLleuLysSerArg 239
 QY 727 GACCTGGGCTGCTGGTTTGACATCAAGTTCACAGCTCTCTATCTGAGCCCTTC 786
 Db 240 AAspLeuGlySerLleLleCysAspLleLysPheSerGluLeuThrTyLLeuAspAlaPhe 259
 QY 787 TGGGGCGACTACTGAGTGGCGCCCTGCTGACAGCCCTGGCGGCGCTTCCTCACTGAG 846
 Db 260 TrpArgAspTrpLleAsnGlySerLeuLeuGlnAlaLeuLysGlyValPheLleThrAsp 279
 QY 847 GCCCTGCGAGAGCTGTGGCGCCGAGGCTGTTGCGCTGCTGTCAGTGTGATGAGCT 906
 Db 280 SerLeuLysGlnAlaValGlyHisGlnLleLleLysLeuLeuValAsnValAspGluGlu 299
 QY 907 GAC 909
 Db 300 Asp 300

RESULT 9

US-09-723-450-6
 ; Sequence 6, Application US/09723450
 ; Patent No. 6576751
 ; GENERAL INFORMATION:
 ; APPLICANT: Alnemrl, Emad S.
 ; TITLE OF INVENTION: Padd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, An

QY 331 TCTCCA-----GAACGTATAGCTATGCACTCCACGCTTTCAAGAGAGAG 381
Db 110 CysProAspLeuValAspLysTyrLeuGlnGluThrSerIleArgTyrValThrProArg 129
QY 382 GGTAGCTGCGCTCCCGCGGAGTCAAGAGTTCTGAAATTCACAGAGGTCAAGTGG 441
Db 130 AlaLeuSerAspProGluProArgProGlnProSerLysThrValProProHisTyr 149
QY 442 GAGACAGGCTCCCGCCCAACCAAG-----CGCAGCGGCGGAGTCCGAGGC 486
Db 150 ProValValCysCysProThrSerGlySerGlnMetCysSerLysArgProAlaArgGly 169
QY 487 CGGCGCAGTGTGTGCGCAGAGCGGCGGAGAGGCGGCCAGCCGACCCACAGCAG 546
Db 170 ArgThrThrLeuGlnGlySerGlnArgLysArgArg----- 180
QY 547 TCAGAGCCCGCCCAACCTTCTCTGAAAGCAAAGTACCTGTGATTCGGCTCCGGT 606
Db 181 ---LysSerValThrProAspProLysGlnLysGlnThrCysAspIleArgLeuArgVal 199
QY 607 CGAGCAGAGTACGTGCGAGCATGGGCGAGCCTTGGAGAGGCGGTGACATCCCGGCGGCC 666
Db 200 ArgAlaGluTyrCysGlnHisGluThrAlaLeuGlnGlnValPheSerAsnLysGln 219
QY 667 CAGCGCTGGCGCGGAGCTGAGCTGTTTGAGGAGCCAGCTGCTGCTGCTGCTGAG 726
Db 220 AspProLeuGlnArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
QY 727 GACCTGGGCTGTGAGTTTGACATCAAGTTTTCAGAGCTCTCTATCTGAGAGCTTC 786
Db 240 AspLeuGlySerIleIleCysAspIleLysPheSerGlnLeuThrTyrLeuAspAlaPhe 259
QY 787 TGGGGCGCATCTAGAGTGGCCCTGCTGAGAGCCGCGGGGCGTTCCTGCTGAG 846
Db 260 TyrArgAspTyrIleAsnGlySerLeuLeuGlnAlaLeuLysGlyValPheIleThrAsp 279
QY 847 GCCCTGCGAGAGCTGTGCGCGGAGGCTGTTCGCTGCTGCTGCTGCTGCTGCTGCT 906
Db 280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGlnGlu 299
QY 907 GAC 909
Db 300 Asp 300
RESULT 11
US-09-733-167A-6
; Sequence 6, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733.167A
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C-DEDD, which is a deletion mutant of human DEDD comprising amino
; OTHER INFORMATION: acids 109-318 of the naturally occurring human DEDD.
US-09-733-167A-6
Alignment Scores:
Pred. No.: 2,396-23 Length: 210
Score: 422.00 Matches: 81

Percent Similarity: 71.61% Conservative: 30
Best Local Similarity: 52.26% Mismatches: 34
Query Match: 24.59% Indels: 10
DB: 4 Gaps: 1
US-10-030-271-1 (1-909) x US-09-733-167A-6 (1-210)
QY 445 ACAGGCTCCCGCCCAACCAAGCGGAGCGGAGTCCGAGCGGCCAGTGTGTGCC 504
Db 48 ThrSerGlyProGlnMetCysSerLysArgProAlaArgAlaThrLeuGlySer 67
QY 505 AGACGCGCGGAGAGAGGCGGCCAGCCGACCCAGCAGCAGTACAGAGCCCGCAACT 564
Db 68 GlnArgLysArgArg-----LysSerValThrPro 77
QY 565 TCCCTGAGCAAGCAAGTACCTGATGATCCGAGTCCGGGTTGAGAGAGTACTGCGAG 624
Db 78 AspProLysGlnLysGlnThrCysAspIleArgLeuThrValArgAlaGluTyrCysGln 97
QY 625 CATGGCGCAGCTTGGAGAGGCGGCGCATCCCGGCGGCCAGCGCTGCGCGCAG 684
Db 98 HisGluThrAlaLeuGlnGlnValAsnValPheSerAsnLysGlnAspProLeuGlnArgGln 117
QY 685 CTGAGCTGTTTGCGCAGGCCAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Db 118 PheGluArgPheAsnGlnAlaAsnThrIleLeuLysSerArgAspLeuGlySerIle 137
QY 745 TGTGACATCAAGTTCTGAGAGCTCTCTATCTGAGAGCCTTGGGCGACTACCTGAGT 804
Db 138 CysAspIleLysPheSerGlnLeuThrTyrLeuAspAlaPheThrAspTyrIleAsn 157
QY 805 GCGCCCTGCTGAGAGCTGTGCGCGGAGGCTGTTCCTGACTGAGAGCCTGCGAGAGCTGTG 864
Db 158 GlySerLeuLeuGlnAlaLeuLysGlyValPheIleThrAspSerLeuLysGlnAlaVal 177
QY 865 GCGCGGAGAGCTGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
Db 178 GlyHisGlnAlaIleLysLeuLeuValAsnValAspGlnGluAsp 192
RESULT 12
US-09-733-167A-5
; Sequence 5, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733.167A
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 114
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: N-DEDD, which is a deletion mutant of human DEDD comprising the a
; OTHER INFORMATION:mino acids 1-114 of the naturally occurring human DEDD.
US-09-733-167A-5
Alignment Scores:
Pred. No.: 6,246-09 Length: 114
Score: 225.50 Matches: 48
Percent Similarity: 66.02% Conservative: 20
Best Local Similarity: 46.60% Mismatches: 32
Query Match: 13.14% Indels: 3
DB: 4 Gaps: 2

FILE REFERENCE:	107196.136		
CURRENT APPLICATION NUMBER:	US/09/252,991A		
CURRENT FILING DATE:	1999-02-18		
PRIOR APPLICATION NUMBER:	US 60/074,788		
PRIOR FILING DATE:	1998-02-18		
PRIOR APPLICATION NUMBER:	US 60/094,190		
PRIOR FILING DATE:	1998-07-27		
NUMBER OF SEQ ID NOS:	33142		
SEQ ID NO:	28921		
LENGTH:	1034		
TYPE:	PRT		
ORGANISM:	Pseudomonas aeruginosa		
US-09-252-991A-28921			

Alignment Scores:			
Pred. NO.:	3,92e-08	Length:	1034
Score:	217.00	Matches:	115
Percent Similarity:	33.93%	Conservative:	18
Best Local Similarity:	29.34%	Mismatches:	128
Query Match:	13.25%	Indels:	131
DB:	4	Gaps:	17

US-10-030-271-1 (1-909) x US-09-252-991A-28921 (1-1034)			
QY	889	CCAGCAGCGGAACAGCCTCCGGCCCAACG	860
Db	2	ProAlaGlyGlyAspProAlaAspSerArgAspAlaCysArgProValAlaAsp	21
QY	859	---CCTTCGCGAGGGCGCTCAGTCAGGAACACGCCCGCA	824
Db	22	ProProGlyAlaGlyProAlaGlnAlaAlaArgProAlaSerHisAlaGlnArgAla	41
QY	823	---GGGCTTCAGCAGGGCC	806
Db	42	AspArgAspArgArgLeuArgArgGlnArgAlaTyrProGlyAlaAlaValGlyArgPro	61
QY	805	---CACTCAGGTAGTGGCCC	788
Db	62	AlaHisArgAlaGlnProAlaArgArgProGluThrPheAspHisAlaArgAlaArgPro	81
QY	787	AGAAGCGGTCCAGATAGAGAGAGCTCTGAGAACTGATGTCACAAACACAGAGCCAGGT	728
Db	82	AlaArgArgProArgArg	90
QY	727	CCCTTGAGCGCAGCAGCTCGCGGTGCGCCCAACAGTCACAGTCCCGCGCA	674
Db	91	ProArgProAlaGlyThrGlyArgProGlyProArgProProAlaAlaAlaAspGlnArg	110
QY	673	---GCGCTTGAGGCGCGCGGAGTCCAGCGCT	638
Db	111	AlaGlnArgThrArgAlaProGlyHisProGlnArgProArgAlaProGlyGlyAlaPro	130
QY	637	AGCGTGGCCAGTCTCGCAGTACTGCTCGCAACCGGAGCC	596
Db	131	ArgArgAlaArgArgArgHisAlaGlnGlyArgSerGlyAlaAlaGlyTyrGlnLeu	150
QY	595	---GSAATGCACAGGTCACTTTGCGCTTCAGAGAAAGCTTGGCGGCT	551
Db	151	ProGlnGlyArgProGlyArgHisProProAlaGlnAlaArgArg	166
QY	550	CTGACTGCTGTGGGGATGCGGTGGGGCCCTCTCGCGCGCGGTTCGACCAACCATGG	491
Db	167	---ArgAlaAlaProLeuArgArgGly---AspArgArgArgLeuProAlaProAlaAlaG	185
QY	490	GCGGCGCCCGGACTCCGCGCGCTGCGCTTGTTGGG	455
Db	185	IuArgProArgProArgGlnProArgProAlaGlyProArgArgHisAlaGlyArgS	205
QY	454	GGAAGCTGTCTCCCACTGACCTGTGTGAGAAATTGCAGAACTGCTGACCGCAGCGG	395
Db	205	erArgGlnArgProAlaAlaProAla	219
QY	394	GACGCGACCTACCTCTGTCTCTTTTGAAGAGCTGAGAGTGCATAGACGTTCTG	335

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 01:51:46 ; Search time 108.904 Seconds

(without alignments)
5517.913 Million cell updates/sec

Title: US-10-030-271-1
Perfect score: 1716
Sequence: 1 atgagcgtacatccgagtcgac.....tcagtgatgagtcgacgac 909

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1401741 seqs, 330541175 residues
Total number of hits satisfying chosen parameters: 2803482

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p model -DEV=x1p
-Q/cgn2/1/US10000271/runat_21032005.153501.26566/app_query.fasta.1.3150
-DB=Published_Applications_AA -QFMT=fastdn -SUFFIX=rapb -MINMATCH=0.1
-LOOPCUT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blisum62
-TRANS=human40.cdi -1ST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10030271 @CGN 1.1 439 @runat_21032005.153501.26566
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	1543	89.9	318	14	US-10-001-254-18	Sequence 18, Appl
2	1543	89.9	326	14	US-10-296-539-1	Sequence 1, Appl
3	1543	89.9	366	14	US-10-106-698-4626	Sequence 4626, Ap
4	1430	83.3	319	9	US-09-925-302-758	Sequence 758, App
5	1430	83.3	319	10	US-09-925-302-758	Sequence 11, Appl
6	1413	82.3	304	14	US-10-013-477-11	Sequence 1220, Ap
7	987	57.5	242	15	US-10-296-115-1220	Sequence 26, Appl
8	889.5	51.8	217	9	US-09-799-777-26	Sequence 4, Appl
9	614.5	35.8	318	9	US-09-935-223-4	Sequence 1, Appl
10	614.5	35.8	318	9	US-09-733-167-1	Sequence 6, Appl
11	608.5	35.5	318	9	US-09-935-223-6	Sequence 3, Appl
12	608.5	35.5	318	9	US-09-733-167-3	Sequence 8, Appl
13	514	30.0	101	14	US-10-001-254-8	Sequence 37, Appl
14	422	24.6	210	9	US-09-733-167-6	Sequence 38, Appl
15	265.5	15.5	146	14	US-10-001-254-38	Sequence 36, Appl
16	227	13.2	99	14	US-10-001-254-36	Sequence 5, Appl
17	227	13.9	19662	15	US-10-084-846A-6	Sequence 11701
18	225.5	13.1	114	9	US-09-733-167-5	Sequence 6, Appl
19	224	13.1	99	14	US-10-001-254-37	Sequence 11701
20	223	13.6	514	16	US-10-437-963-111701	Sequence 6, Appl
21	218	12.7	19662	15	US-10-084-846A-6	Sequence 3, Appl
22	216.5	12.6	19695	15	US-10-084-846A-3	Sequence 8, Appl
23	215.5	12.6	19608	15	US-10-084-846A-8	Sequence 3, Appl
24	214.5	13.1	19695	15	US-10-084-846A-3	Sequence 8, Appl
25	213.5	13.0	19608	15	US-10-084-846A-8	Sequence 14622, A
26	211	12.9	375	14	US-10-156-761-14622	Sequence 3, Appl
27	210	12.8	504	15	US-10-342-331-3	Sequence 165015, A
28	204	12.5	625	16	US-10-437-963-165015	Sequence 10907, A
29	202.5	12.4	797	14	US-10-156-761-10907	Sequence 4, Appl
30	201	12.3	720	15	US-10-342-331-4	Sequence 125004, A
31	200	12.2	19723	15	US-10-084-846A-5	Sequence 7, Appl
32	199.5	12.2	497	16	US-10-437-963-125004	Sequence 14, Appl
33	198	11.5	19652	15	US-10-084-846A-7	Sequence 226, App
34	198	12.1	19652	15	US-10-084-846A-7	Sequence 88, Appl
35	197.5	12.1	1023	10	US-09-893-519A-14	Sequence 72, Appl
36	197	12.0	1472	16	US-10-437-963-136000	Sequence 3, Appl
37	197	11.5	1466	10	US-09-918-715-226	Sequence 10390, A
38	197	11.5	1466	14	US-10-301-822-33	Sequence 145800, A
39	197	11.5	1466	14	US-10-301-822-33	Sequence 141068, A
40	197	11.5	1466	15	US-10-257-021-72	
41	197	11.5	1466	16	US-10-357-851-3	
42	197	11.5	1466	16	US-10-358-024-3	
43	197	11.5	1466	16	US-10-734-563-145800	
44	196.5	12.0	550	16	US-10-437-963-145800	
45	195.5	11.9	390	16	US-10-437-963-141068	

ALIGNMENTS

RESULT 1
US-10-001-254-18
Sequence 18, Application US/10001254
Publication No. US20030049702A1
GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sug Hyung
APPLICANT: Roth, Wilfried
APPLICANT: Stemmer-liewen, Frank
TITLE OF INVENTION: No. US20030049702A1 Death Domain Proteins
FILE REFERENCE: P-LJ 5037
CURRENT APPLICATION NUMBER: US/10/001,254
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FASTSEQ for Windows Version 4.0

; SEQ ID NO 18
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-254-18

Alignment Scores:

Pred. No.:	3.24e-89	Length:	318
Score:	1543.00	Matches:	302
Percent Similarity:	99.67%	Conservative:	0
Best Local Similarity:	99.67%	Mismatches:	1
Query Match:	89.92%	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-1 (1-909) x US-10-001-254-18 (1-318)

```
QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGGAGATGATGCTGACTAC 60
DB 1 MetAlaLeuSerGlySerThrProAlaProCyStrpGlnuAspGluCyLeuAspTyr 20
QY 61 TACGGAGTCTGCTGCTTCAACCGTATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
DB 21 TyGlyMetLeuSerLeuHisArgMetPheGluValGlyGlnLeuThrGluCys 40
QY 121 GACCTGGAGCTCTTGGCTTTTCTGCTGATGAGGCTCTGCGCCGCGGAGGCTTAGCC 180
DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaGlyGlyLeuAla 60
QY 181 CGGGCCCGGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGGCGGCGAGTGGCGAG 240
DB 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuGluLeuGluLeuGluLeu 80
QY 241 AGCAACCTGCGGCTGCTGCGGCAACTCTGCGGCTGCGCCGCGCACTGCTGCGC 300
DB 81 SerAsnLeuArgLeuLeuGluLeuGluLeuLeuAlaArgValLeuAlaArgHisAspLeuPro 100
QY 301 CACCTGGCGGCGCAAGCGGCGCGCGCAGTGTCTTCAGAGCGTATAGCTATGGCACTCC 360
DB 101 HisLeuAlaArgIysArgArgArgProValSerProGluArgIysSerTyrGlyThrSer 120
QY 361 AGCTCTTCAAGAAGGAGAGGAGGAGTGCCTGCGCGCTGCGGCGCAAGCTTGTGCA 420
DB 121 SerSerSerIysArgItrGlnGlySerCysArgArgArgArgItrGlnSerSerSerAla 140
QY 421 AATTCTCAGAGGAGTCACTGAGGAGACAGGCTCCCGCCCAACCAAGCGGCGGAGT 480
DB 141 AsnSerGlnGlnGlnGlnItrpGlnThrGlySerProItrHisArgGlnAlaArgSer 160
QY 481 CGGGCGCGGCGCAGTGTGTGTCGACAGCGCGCGGAGGAGGCGCCGACCGGCA 540
DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgGlyAlaProAlaAlaProGln 180
QY 541 CACGAGTCAAGCGGCGGCGCAACTCTGCTGAGGCGCAAGTGAACCTGTGATCCGGCTC 600
DB 181 GlnGlnSerItrProAlaArgProSerSerGlnGlyValItrCysAspIleArgLeu 200
QY 601 CGGGTTTGAAGAGTACTGAGCATGAGGCGCAGCTTGGAGCGAGCGCTGAGCATCCCG 660
DB 201 ArgValAlaArgIaGluTyrCysGlnHisGlyProAlaLeuGlnGlnGlyValAlaSerArg 220
QY 661 CGGCGCCGAGCGCTGCGGCGGCGAGCTGACGTCGTTGGGCGAGCGCACCGCATGCTCGC 720
DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaItrAlaValLeuArg 240
QY 721 TCAAGGAGCTGGGCTGTGTGTTGTGACATCAAGTTCAGAGCTCTCTATCTGAGAC 780
DB 241 SerArgAspLeuGlySerValValCysAspIleArgHisSerGlnLeuSerTyrLeuAsp 260
QY 781 GCTCTTGGGGCGCACTAGTAGTGGCGCGCTGCTGAGGCGCGCGGCGGCTGTTCTG 840
DB 261 AlaPheItrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuAlaGlyValPheLeu 280
QY 841 ACTGAGGCGCTGCGAGAGCTGTGGGCGGAGGCTGTGCGCTGCTGATGATGAT 900
```

```
DB 281 ThrGluAlaLeuArgGluAlaValGlyArgGluAlaValArgLeuLeuValSerValAsp 300
QY 901 GAGCTGAC 909
DB 301 GluAlaAsp 303
```

RESULT 2

US-10-296-539-1

; Sequence 1, Application US/10296539

; Publication No. US20030165933A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: TANG, Y. Tom

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: YUE, Henry

; APPLICANT: BURFORD, Neil

; APPLICANT: DING, Li

; APPLICANT: ELLIOTT, Vicki S.

; APPLICANT: PATTERSON, Chandra

; APPLICANT: BAUGHN, Mariah R.

; TITLE OF INVENTION: REGULATORS OF APOPTOSIS

; FILE REFERENCE: PI-0307 PCT

; CURRENT APPLICATION NUMBER: US/10/296,539

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: 60/209,407; 60/250,326

; PRIOR FILING DATE: 2000-06-01; 2000-11-30

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PERL Program

; SEQ ID NO 1

; LENGTH: 326

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030165933A1 3102521CD1

US-10-296-539-1

Alignment Scores:

Pred. No.:	3.24e-89	Length:	326
Score:	1543.00 <td>Matches:</td> <td>302</td>	Matches:	302
Percent Similarity:	99.67% <td>Conservative:</td> <td>0</td>	Conservative:	0
Best Local Similarity:	99.67% <td>Mismatches:</td> <td>1</td>	Mismatches:	1
Query Match:	89.92% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-1 (1-909) x US-10-296-539-1 (1-326)

```
QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGGAGATGATGCTGACTAC 60
DB 1 MetAlaLeuSerGlySerThrProAlaProCyStrpGlnuAspGluCyLeuAspTyr 20
QY 61 TACGGAGTCTGCTGCTTCAACCGTATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
DB 21 TyGlyMetLeuSerLeuHisArgMetPheGluValGlyGlnLeuThrGluCys 40
QY 121 GACCTGGAGCTCTTGGCTTTTCTGCTGATGAGGCTCTGCGCCGCGGAGGCTTAGCC 180
DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaGlyGlyLeuAla 60
QY 181 CGGGCCCGGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGGCGGCGAGTGGCGAG 240
DB 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuGluLeuGluLeuGluLeu 80
QY 241 AGCAACCTGCGGCTGCTGCGGCAACTCTGCGGCTGCGCCGCGCACTGCTGCGC 300
DB 81 SerAsnLeuArgLeuLeuGluLeuGluLeuLeuAlaArgValLeuAlaArgHisAspLeuPro 100
QY 301 CACCTGGCGGCGCAAGCGGCGCGCGCAGTGTCTTCAGAGCGTATAGCTATGGCACTCC 360
DB 101 HisLeuAlaArgIysArgArgArgProValSerProGluArgIysSerTyrGlyThrSer 120
QY 361 AGCTCTTCAAGAAGGAGAGGAGTGAAGTGCCTGCGCGCTGCGGCGAGTCAAGCTTCTGCA 420
```

```

Db      121 SerSerSerlysaGrThrGlnGlySerCyseArGrArGrArGrGlnSerSerSerSerAla 140
Qy      421 AATTCTCAGACGAGGCTGATGGAGACAGAGGCTCCCCCAACCAAGCGGACGCGCGAGT 480
Db      141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
Qy      481 CGGGCCCGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db      161 ArgGlyArgProSerArgGlyAlaArgArgArgArgArgArgArgArgArgArgArgArg 180
Qy      541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAGAGCAAGTGAACCTGTGACATCCGCTC 600
Db      181 GlnGlnSerGlnProAlaArgProSerSerSerGlnGlySerValThrCyseAspIleArgLeu 200
Qy      601 CGGGTTGACAGAGTCTGCGAGCATGGGCGACGCTTGAGCGGGGCGGCGCTGCGCATCCCG 660
Db      201 ArgValArgAlaGlnArgCyseGlnHisGlyProAlaLeuGlnGlnGlnGlnGlnGlnGln 220
Qy      661 CGGCCCCGAGCGCTGCGCGCGGCGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db      221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
Qy      721 TCAAGGACCTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db      241 SerArgAspLeuGlySerValValCyseAspIleIleYsPheSerGlnLeuSerTyLeuAsp 260
Qy      781 GCCTTCGGGGCGCATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db      261 AlaPheThrGlyAspTyLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
Qy      841 ACTGAGCCCTGCGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db      281 ThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerValAsp 300
Qy      901 GAGGCTGAC 909
Db      301 GlnAlaAsp 303

```

RESULT 3

```

US-10-106-698-4626
; Sequence 4626, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4626
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-4626

```

Alignment Scores:

```

Pred. No.: 3,25e-89 Length: 366
Score: 1543.00 Matches: 302
Percent Similarity: 99.67% Conservative: 0
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 89.92% Indels: 0
DB: 14 Gaps: 0

```

US-10-030-271-1 (1-909) x US-10-106-698-4626 (1-366)

```

Qy      1 ATGGCGTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGATGCTCTGAGTAC 60
Db      41 MetAlaLeuSerGlySerThrProAlaProCyseThrPbGlnAspGluCyseLeuAspTy 60
Qy      61 TACGGATGCTGTTCGCTTCAACCGTATGTTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Db      61 TyrlYMetLeuSerLeuHisArgMetPheGlnValAlaGlyGlyGlnLeuThrGlnCys 80
Qy      121 GAGTGGAGCTCTGAGCTTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Db      81 GlnLeuGlnLeuLeuAlaPheLeuLeuAspGlnAlaProGlyAlaAlaGlyGlyLeuAla 100
Qy      181 CGGGCCCGGCGCAGCTTGAAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db      101 ArgAlaArgSerGlyLeuGlnLeuLeuGlnLeuGlnLeuGlnLeuGlnLeuGlnLeuGln 120
Qy      241 AGCAACTGGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db      121 SerAsnLeuArgLeuLeuGlnGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 140
Qy      301 CACCTGCGCGCAGCGCGCGCGCGAGTGTCTCAGAGCTTATGACCTATGACCTTC 360
Db      141 HisLeuAlaArgGlySerArgArgArgProValSerProGlnArgTySerTyrlYThrSer 160
Qy      361 AGCTTTCAAGAAGACAGAGGAGTACCTGCGCGCGCGAGTCAAGCAAGTGTGTGTGTGTGTGT 420
Db      161 SerSerSerlysaGrThrGlnGlySerCyseArGrArGrArGrGlnSerSerSerAla 180
Qy      421 AATTCTCAGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db      181 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Qy      481 CGGGCCCGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db      201 ArgGlyArgProSerArgGlyAlaArgArgArgArgArgArgArgArgArgArgArg 220
Qy      541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAGAGCAAGTGAACCTGTGACATCCGCTC 600
Db      221 GlnGlnSerGlnProAlaArgProSerSerSerGlnGlySerValThrCyseAspIleArgLeu 240
Qy      601 CGGGTTGACAGAGTCTGCGAGCATGGGCGACGCTTGAGAGCGGCGCTGTGACATCCCGG 660
Db      241 ArgValArgAlaGlnArgCyseGlnHisGlyProAlaLeuGlnGlnGlnGlnGlnGlnGln 260
Qy      661 CGGCCCCGAGCGCTGCGCGCGGCGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db      261 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 280
Qy      721 TCAAGGACCTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db      281 SerArgAspLeuGlySerValValCyseAspIleIleYsPheSerGlnLeuSerTyLeuAsp 300
Qy      781 GCCTTCGGGGCGCATCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db      301 AlaPheThrGlyAspTyLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 320
Qy      841 ACTGAGCCCTGCGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db      321 ThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerValAsp 340
Qy      901 GAGGCTGAC 909
Db      341 GlnAlaAsp 343

```

RESULT 4

```

US-09-925-302-758
; Sequence 758, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302

```

```

; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PR
; ORGANISM: Homo sapiens
; US-09-925-302-758

Alignment Scores:
Pred. No.: 4,31e-82 Length: 319
Score: 1430.00 Matches: 278
Percent Similarity: 99.29% Conservative: 0
Best Local Similarity: 99.29% Mismatches: 1
Query Match: 83.33% Indels: 1
DB: Gaps: 0

US-10-030-271-1 (1-909) x US-09-925-302-758 (1-319)
QY 1 ATGGCGCTATCCGGGTGACCCCGCGCTGCTGGAGAGAGATGATGCTGACTAC 60
DB 41 MetAlaLeuSerGlySerThrProAlaProCyStrpGluGluaspGluCysLeuAspTyr 60
QY 61 TACGGGATGCTGCTTACCGTATGTTGAGATGGTGGCGCGCACTGACCGAGTGC 120
DB 61 TyrglyMetLeuSerLeuHisArgMetPheGluValAlaGlyGlnLeuThrGluCys 80
QY 121 GAGCTGAGAGCTCTGCGCTTTCTGCTGATGAGGCTCTCGCGCGCGGAGGCTTACC 180
DB 81 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 100
QY 181 CGGGCCCGACGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGGCGAG 240
DB 101 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuGluLeuGluGlnCysAspGlu 120
QY 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGGCTGCTGCGCGCGCACTGCTGCCG 300
DB 121 SerAsnLeuArgLeuLeuGluGlnLeuLeuValLeuAlaArgHisAspLeuLeuPro 140
QY 301 CACCTGGCGCGCAAGCGCGCGCGCGCACTGCTTCCAGACGCTATAGCACTGCC 360
DB 141 HisLeuAlaArgGlySerArgArgArgProValSerProGluArgGlySerGlyThrSer 160
QY 361 AGCTTTCAAGAGGACAGAGGAGTGCCTGCGCGCTGCGCGAGTCAACAGCTTGCA 420
DB 161 SerSerSerGlyArgThrGlnGlySerCysArgArgArgGlnSerSerSerSerAla 180
QY 421 AATTCTCAGCAGGAGTCACTGAGAGACAGGCTCCCCCAACCAAGCGCGAGGAGT 480
DB 181 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
QY 481 CGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 540
DB 201 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 220
QY 541 CACGAGTCAAGCGCGCGCGCACTTCTGAGAGCGCAAGTGAACCTGATCCGCGCTC 600
DB 221 GlnGlnSerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeu 240
QY 601 CGGGTTCAAGCAGATCACTGAGCATGAGCGCGCGCTTGGAGAGAGGCGCTGCGATCCCG 660
DB 241 ArgValArgAlaGluTyrCysGlnHisGlyProAlaLeuGluGlnGlnGlnValAlaSerArg 260
QY 661 CGGCGCGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGC 720
DB 261 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 280
QY 721 TCAAGGAGCTGGGCTTGTGGTTTGTGACATCAAGTCTTCAAGGCTCTCTATCTGGAC 780

```

```

DB 281 SerArgAspLeuGlySerValValCysAspIleLeuPheSerGluLeuSerTyrLeuAsp 300
QY 781 GCCTTCTGGCGGCACTACTGAGTGCGCGCTGCTGCGAGCCCTGCGGGCGCTTCC 838
DB 301 AlaPheTrpGlyAspPyrLeuSerGlyAlaLeuGln-ProCysGlyAlaCysSer 319

RESULT 5
US-09-925-302-758
; Sequence 758, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PR
; ORGANISM: Homo sapiens
; US-09-925-302-758

Alignment Scores:
Pred. No.: 4,31e-82 Length: 319
Score: 1430.00 Matches: 278
Percent Similarity: 99.29% Conservative: 0
Best Local Similarity: 99.29% Mismatches: 1
Query Match: 83.33% Indels: 1
DB: Gaps: 0

US-10-030-271-1 (1-909) x US-09-925-302-758 (1-319)
QY 1 ATGGCGCTATCCGGGTGACCCCGCGCTGCTGGAGAGAGATGATGCTGACTAC 60
DB 41 MetAlaLeuSerGlySerThrProAlaProCyStrpGluGluaspGluCysLeuAspTyr 60
QY 61 TACGGGATGCTGCTTACCGTATGTTGAGATGGTGGCGCGCACTGACCGAGTGC 120
DB 61 TyrglyMetLeuSerLeuHisArgMetPheGluValAlaGlyGlnLeuThrGluCys 80
QY 121 GAGCTGAGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTACC 180
DB 81 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 100
QY 181 CGGGCCCGACGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGCGAGTGGCGAG 240
DB 101 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuGluLeuGluGlnCysAspGlu 120
QY 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGGCTGCTGCGCGCGCACTGCTGCCG 300
DB 121 SerAsnLeuArgLeuLeuGlnLeuLeuValLeuAlaArgHisAspLeuLeuPro 140
QY 301 CACCTGGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGTGC 360
DB 141 HisLeuAlaArgGlySerArgArgArgProValSerProGluArgGlySerGlyThrSer 160
QY 361 AGCTTTCAAGAGGACAGAGGAGTGCCTGCGCGCTGCGCGAGTCAACAGCTTGCA 420
DB 161 SerSerSerGlyArgThrGlnGlySerCysArgArgArgGlnSerSerSerSerAla 180
QY 421 AATTCTCAGCAGGAGTCACTGAGAGACAGGCTCCCCCAACCAAGCGCGAGGAGT 480
DB 181 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
QY 481 CGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 540
DB 201 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 220

```


QY	541	TAGCAGTCAAGACC	CCGCCAGACCTTC	CTGAAAGGAATGAC	CTGTGATCATCCGAGT	600
Db	221	GtngInserGlpIroAlaArgProSerSeegUgLyIyValThrCysApIleArgLeu	240			
QY	601	CGGgTTCGACGAGTACTCGAGCAGTGGGCCAGCCTTGAGAGCAGGGCGTGGCATCCGG	660			
Db	241	ArgValArgAlaGILuTyrCyseGluHISgILyProAlaLeuGluGInGlyAlaIAsaSerArg	260			
QY	661	CGGGCCCGAGGCGCTGGCGCGCGCAGCGTGAACGTGTTGGGCAAGGCCACCGAGAGCGTGGCGC	720			
Db	261	ArgProGlnAlaLeuAlaArgGInLeuAspValPheGILyGlnAlaThrAlaValLeuArg	280			
QY	721	TCAAGGACCTGGGCGCTGTGTGTGTGTGACATCAATGTTCTCAGAGCGTCTCTCATTCGAC	780			
Db	281	SerArgAspLeuGILySerValValCysAspIleILyPheSerGILyLeuSerTyrLeuAsp	300			
QY	781	GCCTTCTGGGGGACATCACTGAGTGGCGGCGCTGTCAGAGCCCTGGGGGCGGTGTTCC	838			
Db	301	AlaPheTyrPGLyAspTyrLeuSerGILyAlaLeuLeuGln-ProCysGILyAlaCysSer	319			

RESULT 6

```

US-10-013-477-11
; Sequence 11, Application US/10013477
; Publication No. US20030049732A1
; GENERAL INFORMATION:
; APPLICANT: N1 et al.
; TITLE OF INVENTION: Apoptosis Related Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT002P1
; CURRENT APPLICATION NUMBER: US/10/013,477
; CURRENT FILING DATE: 2001-12-13
; PRIOR APPLICATION NUMBER: 09/669,445
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: PCT/US00/06642
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/126,018
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/139,638
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/149,449
; PRIOR FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 11
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-013-477-11

```

Alignment Scores:

Pred. No.:	5.07e-81	Length:	3
Score:	1413.00	Matches:	2
Percent Similarity:	99.64	Conservative:	0
Best Local Similarity:	99.64	Mismatches:	1
Query Match:	82.348	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-1 (1-909) x US-10-013-477-11 (1-304)

QY	67	ATGCTGTCCGCTTACCCGTAATGTTCCAGAGTGTGGCCGAGCACTGAACCGAGTCCGAGCTG	126
DB	1	MetLeuSerLeuH1aArgMetCpHeGluValGlyGlyGluLeuThrGluCysGGluLeu	20
QY	127	GAGCTCTCGGCTTTCTGCTGCAATGAAGGCTCTGCGCGCGCCGAGAGCTTTAGCCCGAGCC	186
DB	21	GluLeuLeuAlaPheLeuLeuAspGluIaIaProGlyAlaIlaGlyLeuAlaIaArgIa	40
QY	187	CCGAGCGGCTTGAAGCTCTGCTGCAAGCTTGAAGCGCGCGGAGCAATGTGGCGGAGAGCAAC	246
DB	41	ArgSerGlyLeuGluLeuLeuLeuGluLeuGluIaArgGlyIaGlyCysAspGluSerAsn	60
QY	247	CTGCGGCTCTGGGGGCAACTCTGCGCGGTCTGGCCGCGCAAGCACTGTGCGCGCACTG	306

Db	61	LeuArgLeuLeuGlyGlnLeuLeuArgAlaLeuAlaArgHisAspLeuLeuProHisIleu	80
Oy	307	GGCGGCAGAGCGGCGCGGCGAGGTGTCTCCAGAACGCTATAGTCACCTTCAGCTCT	366
Db	81	AlaArgIleValArgArgArgProValSerProGluArgTyrSerTyrGlyHisSerSer	100
Oy	367	TCAAAGAGACAGAGAGGTAGGCTGGCGCGCTCGGAGACTCAAGCAGTTCCTGCAATTCT	426
Db	101	SerIysArgHisGlnGlnGlySerCysArgArgArgGlnIleSerSerSerAlaIleSer	120
Oy	427	CAGCAGGGTCAGTGTGGAGACAGAGCTCCCCCGCCACCAAGCGGACGGCGAGAGTCGGGC	486
Db	121	GlnGlnGlyGlnTrrGlnIleThrGlySerProProTrpIleArgGlnArgArgSerArgGly	140
Oy	487	CGGCGCCATGCTGTGTCCAGAACGGGGGGGAGAGAGGGCGCCACCGGACCCAGACAGCAG	546
Db	141	ArgProSerGlyGlyAlaArgArgArgArgArgGlyAlaProAlaIleProGlnGlnGln	160
Oy	547	TCAGAGCGCGGCGAGACTTCTCTCTTCAAGGCAAGTAGACTGTGACATCCGGCTCCGGATT	606
Db	161	SerGlnProAlaArgProSerSerSerGlnGlyIleValIleHisCysAspIleArgLeuArgVal	180
Oy	607	CGAGCAGAGTACTGCGAGCATGCGGCGAGCTTTGGAGACAGGGCGGTGGCATTCGCGCGGCC	666
Db	181	ArgAlaGlnTyrCysGlnHisIleGlyProAlaLeuGlnGlnGlnIleValAlaIleSerArgArgPro	200
Oy	667	CAGGCGCTGGGGCGGCGGAGCTGGACGTGTTGGGAGAGGCGACGCGAGTGGCGGCTCAAG	726
Db	201	GlnAlaIleValAlaArgGlnIleuAspValPheGlyGlnAlaIleThrAlaValaLeuArgSerArg	220
Oy	727	GACCTGGGCTCTGTGGTGTGTGACATCAAGTTCCTAGAGCTCTCTATCTGAGCGCTTC	786
Db	221	AspLeuGlySerValValaCysAspIleIleHisPheSerGlnLeuSerTyrLeuAspAlaPhe	240
Oy	787	TGGGGCGACTACCTGAGTGGCGGCGCTGCGTGAAGGCGCTGGCGGCGGTTCCTGACATGAG	846
Db	241	TrrGlyAspTyrLeuSerSerGlyAlaLeuLeuGlnAlaLeuArgGlyAlaPheLeuTrpGln	260
Oy	847	GCCCTGCAGAGAGGCTGTGTGGCGCGGAGAGGCTGTTCCGCTTGCCTGTCAGTGTGATGAGGCT	906
Db	261	AlaIleuArgGlnAlaValaGlyArgGlnAlaValaArgLeuLeuValSerValaAspGlnAla	280
Oy	907	GAC 909	
Db	281	Asp 281	

RESULT 7

US-10-296-115-1220
; Sequence 1220, Application US/10296115
; Publication No. US20040053248A1

APPLICANT: Hysed In

```

1  TITLE OF INVENTION: NO. US20040053248A1 Nucleic Acids and Polypeptides
2
3  FILE REFERENCE: 784PCT
4
5  CURRENT APPLICATION NUMBER: US/10/296,115
6
7  CURRENT FILING DATE: 2002-11-18
8
9  PRIOR APPLICATION NUMBER: US09/468,725
10
11 PRIOR FILING DATE: 2000-01-21
12

```

PRIOR FILING DATE: 2000-04

```

; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1220
; LENGTH: 242
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-296-115-1220

```

Alignment Scores:

Pred. No.:	3.61e-54	Length:	1
Score:	987.00	Matches:	2
Percent Similarity:	97.52%	Conservative:	1
Best Local Similarity:	97.03%	Mismatches:	2
Query Match:	57.52%	Indels:	3

DB: 15 Gaps: 1

US-10-030-271-1 (1-909) x US-10-296-115-1220 (1-242)

QY 310 CGCAAGCGGGCGGGCA-----GTGTCACGAAGCTATAGCTATGACCTCCAGC 363

DB 18 ArgSerHisArgThrProGlnSerValSerProGlnArgThrGlyThrSerSer 37

QY 364 TCTTCAAGAGACAGAGGGTAGCTGCGCTGCGGAGTCAAGCACTTTCGCAAT 423

DB 38 SerSerLysArgThrGlnGlySerCysArgArgArgArgLysSerSerSerAlaAsn 57

QY 424 TCTCAGCAGGGGTAGTGGGAGACAGGTCTCCCGCCCAACAGCGGCGCGGAGTCCG 483

DB 58 SerGlnGlnGlyGlnThrGlnThrGlySerProThrLysArgGlnArgSerArg 77

QY 484 GGCAGCGCCGAGTGGTGGTCAGACGGCGGAGAGAGGGCCCGAGCGCCAGCCCGACG 543

DB 78 GlyArgProSerLysGlyAlaLysArgArgArgGlyAlaProAlaAlaProGlnGln 97

QY 544 CAGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAAGTGAATCCGGCTCCGG 603

DB 98 GlnSerGlnProAlaArgProSerSerGlnGlyLysValThrCysAspLysLeuArg 117

QY 604 GTTCAGCAGAGTACTGCGCAGCATGGGCGCAGCTTGGAGCAGGGCGGTGCGATCCGGCGG 663

DB 118 ValArgAlaGlnLysCysGlnHisGlyProAlaLeuGlnGlnGlyAlaAlaSerArgArg 137

QY 664 CCCAGCGCGTGGCGGGCGGAGCTGGAAGTGGGAGGCGCAAGGAGTGGCGCTGCA 723

DB 138 ProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArgSer 157

QY 724 AGGAGCCTGGGCTCTGTGTTTGTGACATCAAGTTCACAGAGCTCTCTTATCTGAGAGCC 783

DB 158 ArgAspLeuLysSerValValCysAspLysPheSerGlnLeuSerThrLeuAspAla 177

QY 784 TTTCTGGGGCAGCACTCAAGTGGCGCGCTCTGTGAGGCGCTGGGGCGGTCTGACT 843

DB 178 PheTrpGlyAspTrpLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeuThr 197

QY 844 GAGGCGCTGGAGAGGCTGTGGGCGGGGAGGAGTGTGGCTGCTGCTGCTGAGTGGAG 903

DB 198 GlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerValAspGln 217

QY 904 GCTGAC 909

DB 218 AlaAsp 219

RESULT 8

US-09-799-777-26

Sequence 26, Application US/09799777

Patent No. US20020091244A1

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

Hillman, Jennifer L.

Corley, Neil C.

Guegler, Karl J.

Baugh, Mariah

Sather, Susan

Shah, Purvi

TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS

NUMBER OF SEQUENCES: 154

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/799, 777

FILING DATE: 06-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/002,485

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BILLINGS, LUCY J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0459 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 217 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: UTRAMOT06

CLONE: 1638407

SEQUENCE DESCRIPTION: SEQ ID NO: 26 :

US-09-799-777-26

Alignment Scores:

Pred. No.:	Length:	217
Score:	889.50	183
Percent Similarity:	94.33%	0
Best Local Similarity:	94.33%	8
Query Match:	51.84%	1
DB:	Gaps:	1

US-10-030-271-1 (1-909) x US-09-799-777-26 (1-217)

QY 350 ATGGCACTTCACCTTTCAAGAGAGACAGAGGTTAGCTGCGCTGCGGAGTCAA 409

DB 1 MetAlaProProAlaLeuGlnArgGlyGlnArgValAlaAlaValAlaValGlySerGln 20

QY 410 GCAGTCTGCAATTTCTCAGCAGGAGTCAAGTGGAGACAGGCTCCCGCCCAACAGCGGC 469

DB 21 AlaValLeuGlnLysLeuSerArgValSerGlyArgGlnAlaProProGlnProSerGly 40

QY 470 AGCGGCGAGTCCGGGCGCGCCAGTGTGTGTCAGACGCGCGCGAGAGGGGCGCCAG 529

DB 41 SerGlyGlyValGlyAlaGlyProValValProAspGlyGlyGlyGlyProGln 60

QY 530 CGGACCCCGACAGCAGTCAAGACCCCGCCAGACCTTCTCTGAAGCMAA----- 579

DB 61 ProHisProSerSerSerGlnSerProProAspLeuProLeuLysAlaGlyAspThrVal 80

QY 580 -----GTGACCTGTGACATCCGGCTCCGGGTTTCAGACAGTACTGCGAGCAT 627

DB 81 MetGlyLysGlnAlaGlnArgAspLysArgLeuArgValArgAlaGlnLysGlnHis 100

QY 628 GGGCGACCTTGGAGAGAGGCGTGGCATCCCGCGGCGCCCGAGCGCTGGCGCGCAGCTG 687

DB 101 GlnProAlaLeuGlnGlnGlyValAlaSerArgArgProGlnAlaLeuAlaArgGlnLeu 120

QY 688 GAGCTGTTGGGAGGCGCCAGCGAGTGTGCTCAAGGAGCCTGGGCTGTGTGTTGT 747

DB 121 AspValPheGlyGlnAlaThrAlaValAlaLeuArgSerArgAspLeuGlySerValCys 140

QY 748 GACATCAAGTTCACAGCTCTCTTCTGAGAGCTTCTGGGGGACATCACTGAGTGGC 807

DB 141 AspLysPheSerSerGlnLeuSerThrLeuAspAlaPheTrpLysArgLysSerGly 160

QY 808 GCGCTGTGAGAGCGCTGGCGGCGGTGCTCTACTGAGGCGCTGAGAGGCTGTGGGC 867

DB 161 AlaLeuGlnAlaLeuArgGlyValPheLeuThrGlnAlaLeuArgGlnAlaValGly 180


```

QY 151 GAGGCTCTGGCCGCCGCGGAGCTTACCCCGGCGCGGAGCGGCTTACGCTCTG 210
Db 52 Val11leAspAspHisg1uArgg1yLeu-----11leArgAsn1yAspPheLeuLeu 69
QY 211 GAGCTGAGAGCGCGCGGAGCTGCGGCGGAGCAACTTGGCGCTGAGGCAACTCTCTG 270
Db 70 AlaleuGlnArgGlnArgCysAspGlnSerAsnHeArgGlnVal11leuGlnLeuLeu 89
QY 271 CGGCTGTGCGCGCGGAGCTGCTGCTGCGGAGCTGCGGCGGAGCGGCGCGGAGCTG 330
Db 90 Arg11leThrArgHisAspLeuLeuProTyrVal11ThrLeuYsArgArgAlaVal 109
QY 331 TCTCCA-----GACCGCTAAGCTATGCGACTCC-----AGCTCTCAAG 372
Db 110 CysProAspLeuValAlaAsp1yTyrLeuGlnGln1yHisSer11leArgTyrVal11ThrProArg 129
QY 373 AGGACAGAGGAGTACCTCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAG 432
Db 130 AlaleuSerAspProGlnProArgProArgProArgProArgProArgProArgProArgPro 149
QY 433 GGTGAG-----TGGAGAGAGGCTCCCGGCAAGCGGAGCGGAGCGGAGCTGCGGAG 486
Db 150 ProValValCysCysProThrSerGlyProGlnMetCysSer1yAspProAlaArgGly 169
QY 487 CGGCGGAGTGTGTGTGCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
Db 170 ArgAla11ThrLeuGlnArg1ySerGlnArg1yAspArg-----180
QY 547 TCAGAGCGCGGAGAGCTTCTGTAAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 606
Db 181 ---LysSerVal11ThrProAspProArgProArgProArgProArgProArgProArgPro 199
QY 607 CGAGCAGAGTACTGCGAGCATGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 666
Db 200 ArgAlaGln1yTyrCysGln1yHisg1uThrAlaLeuGln1yAsnVal11PheSerAsn1yGln 219
QY 667 CAGCGCTGCGCGCGGAGCTGCGGAGCTGTTGGAGAGCGGAGCGGAGCTGCGGAGCGGAG 726
Db 220 AspProLeuGln1yArgGlnPheGln1yArgPheAsnGln1yAsnThr11leu1ySerArg 239
QY 727 GACCTGGAGCTCTGTGTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 786
Db 240 AspLeuGlnSer11le1yCysAsp11le1yPheSerGlnLeu1yTyr11leuAspAlaPhe 259
QY 787 TGGGCGCACTAAGCTGAGCGGCGGAGCTGCTGCGGAGCGGAGCGGAGCTGCTGCTGAG 846
Db 260 TrpArgAsp1yTyr11leuGln1ySer11leuGln1yAlaLeu1yGln1yVal11Phe11leThrAsp 279
QY 847 GCCTGCGAGAGGCTGCGGCGGAGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
Db 280 Ser11leuYsGln1yAlaVal11Gln1yHisg1uAla11le1ySerLeuLeuVal11AspGln 299
QY 907 GAC 909
Db 300 Asp 300

```

```

; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
; US-09-935-223-6

Alignment Scores:
Pred. No.: 2,66e-30 Length: 318
Score: 608.50 Matches: 131
Percent Similarity: 62.13% Conservative: 56
Best Local Similarity: 43.52% Mismatches: 93
Query Match: 35.46% Indels: 21
DB: Gaps: 5

US-10-030-271-1 (1-909) x US-09-935-223-6 (1-318)

QY 34 TGGAGAGGAGTATGTCCTGAGCTACGAGATGCTGTGCTTACCGCTATGTTGAG 93
Db 12 TrpProGlnGln1yArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 31
QY 94 GTGCTGCGCGGAGCACTGACCGAGTGCAGCTGAGCTGCTGCGCTTCTG-----CTGGAT 150
Db 32 11leValG1yThr11leuThrHisAspArgVal11ArgVal11leuSerPheLeuValAsp 51
QY 151 GAGGCTCTGGCCGCCGCGGAGCTTACCCCGGCGCGGAGCGGAGCTTACGCTCTG 210
Db 52 Val11leAspAspHisg1uArgg1yLeu-----11leArgAsn1yAspPheLeuLeu 69
QY 211 GAGCTGAGAGCGCGCGGAGCTGCGGAGCAACTTGGCGCTGAGGCAACTCTCTG 270
Db 70 AlaleuGlnArgGlnArgCysAspGlnSerAsnHeArgGlnVal11leuGlnLeuLeu 89
QY 271 CGGCTGTGCGCGCGGAGCTGCTGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTG 330
Db 90 Arg11leThrArgHisAspLeuLeuProTyrVal11ThrLeuYsArgArgAlaVal 109
QY 331 TCTCCA-----GACCGCTAAGCTATGCGACTCC-----AGCTCTCAAG 372
Db 110 CysProAspLeuValAlaAsp1yTyrLeuGlnGln1yHisSer11leArgTyrVal11ThrProArg 129
QY 373 AGGACAGAGGAGTACCTCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAG 432
Db 130 AlaleuSerAspProGlnProArgProArgProArgProArgProArgProArgProArgPro 149
QY 433 GGTGAG-----TGGAGAGAGGCTCCCGGCAAGCGGAGCGGAGCGGAGCTGCGGAG 486
Db 150 ProValValCysCysProThrSerGlyProGlnMetCysSer1yAspProAlaArgGly 169
QY 487 CGGCGGAGTGTGTGTGCGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 546
Db 170 ArgAla11ThrLeuGlnArg1ySerGlnArg1yAspArg-----180
QY 547 TCAGAGCGCGGAGAGCTTCTGTAAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 606
Db 181 ---LysSerVal11ThrProAspProArgProArgProArgProArgProArgProArgPro 199
QY 607 CGAGCAGAGTACTGCGAGCATGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 666
Db 200 ArgAlaGln1yTyrCysGln1yHisg1uThrAlaLeuGln1yAsnVal11PheSerAsn1yGln 219
QY 667 CAGCGCTGCGCGCGGAGCTGCGGAGCTGTTGGAGAGCGGAGCGGAGCTGCGGAGCGGAG 726
Db 220 AspProLeuGln1yArgGlnPheGln1yArgPheAsnGln1yAsnThr11leu1ySerArg 239
QY 727 GACCTGGAGCTCTGTGTTTGAACATGATGATGATGATGATGATGATGATGATGATGATGAT 786
Db 240 AspLeuGlnSer11le1yCysAsp11le1yPheSerGlnLeu1yTyr11leuAspAlaPhe 259
QY 787 TGGGCGCACTAAGCTGAGCGGCGGAGCTGCTGCGGAGCGGAGCTGCTGCTGAG 846

```

```

Db      260 TPrpArgAspTyrIleAsnGlySerLeuLeuGluAlaLeuLysGlyValPheIleThrAsp 279
Qy      847 GCCCTGCGAGAGCTGTGGCGCGGAGGCTGTGCTGCTGTCAGTGGATGAGAGCT 906
Db      280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuValAsnValAspGlu 299
Qy      907 GAC 909
Db      300 Asp 300

RESULT 12
US-09-733-167-3
; Sequence 3, Application US/09733167
; Patent No. US20020099009A1
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT FILING DATE: US/09/733,167
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US20020099009A1 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-167-3

Alignment Scores:
Pred. No.: 2,666-30 Length: 318
Score: 608.50 Matches: 131
Percent Similarity: 62.13% Conservative: 56
Best Local Similarity: 43.52% Mismatches: 93
Query Match: 35.46% Indels: 21
DB: 9 Gaps: 5

US-10-030-271-1 (1-909) x US-09-733-167-3 (1-318)
Qy      34 TGGAGAGAGATGATGCTGCTGAGACTAGCGGAGATGCTGTGCTTACCGTATGTTGAG 93
Db      12 TTrpArgGluGluValGlyGluGlnGluHisGlyLeuTyrSerLeuHisArgPheAsp 31
Qy      94 GTGGTGGCGCGGCAACTGACCGAGTGCAGCTGAGCTGAGCTCTGCTTCTG---CTGGAT 150
Db      32 lIleValGlyThrHisLeuThrHisArgAspValArgValLeuSerPheLeuPheValAsp 51
Qy      151 GAGGCTCTGGCGCGCGCGGAGGCTTAGCGCGCGCGCGCGAGCGGCTTAGAGCTCTGCTG 210
Db      52 ValIleAspAspHisGluArgGlyLeu-----lIleArgAsnGlyArgAspPheLeuLeu 69
Qy      211 GAGCTGAGCGCGCGCGGAGGAGGCGGAGGAGCAACCTGCGCTGCTGGGCACTCTG 270
Db      70 AlalaLeuGluArgGlnGlyArgCysAspArgIleSerAsnPheArgGlnValLeuGlnLeuLeu 89
Qy      271 CGCGTGTGCGCGCGCGCAAGACTGTGCTGCGCACTGCGCGCGCAAGCGCGCGCACTG 330
Db      90 ArgIleIleThrArgHisAspLeuLeuProTyrValThrLeuLysValArgAlaVal 109
Qy      331 TCTGCA-----GAACGCTATAGCTATGCGACTCTTCAAGTCTTCAAAAGACAGAG 381
Db      110 CysArgProAspLeuValAspLysTyrLeuGlnLysIleArgTyrValThrProArg 129
Qy      382 GGTAGCTCGCGTGGCGCGTGGAGTCAAGCAAGTTCGCAAAATTCACAGAGGTCAGTGG 441
Db      130 AlaLeuSerAspProGluProArgProArgProGlnProSerLysThrValProProHisTyr 149
Qy      442 GAGACAGGCTCCCCCAACCAAG-----CGGCAAGCGCGGAGTCTGGGGC 486
```

```

Db      150 ProValValCysCysProThrSerGlySerGlnMetCysSerLysArgProAlaArgGly 169
Qy      487 CGGCGCCAGTGTGTGCGCAAGCGGCGGAGAGAGAGGCGCCAGCGGCAACCCACAGAGAG 546
Db      170 ArgThrThrLeuGlySerGlnArgLysArgArg----- 180
Qy      547 TCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAACCTGTGACATCCGAGTCCGAGTT 606
Db      181 ---LysSerValThrProAspProLysGlnLysGlnThrCysAspIleArgLeuArgVal 199
Qy      607 CGAGCAAGATGACTCGAGACATGAGCCAGCTTGAACAGAGCGTGGCATCCCGCGGCC 666
Db      200 ArgAlaGluTyrCysGlnHisGluThrAlaLeuGlnGlnLysAsnValPheSerAsnLysGln 219
Qy      667 CAGCGCTGGCGCGCGCGCTGAGAGTGTGGGAGGCGACCGGACGAGTCTGCGTCAAG 726
Db      220 AspProLeuGluArgGlnPheGlnArgPheAsnGlnAlaAsnThrIleLeuLysSerArg 239
Qy      727 GACCTGGGCTCTGTGTTGTGACATGAAGTTCAGAGCTCTGCTATCTGAGCGCTTC 786
Db      240 AspLeuGlySerIleIleCysAspIleLysPheSerGluLeuThrTyrLeuAspAlaPhe 259
Qy      787 TGGGCGACTGACTGAGTGGCGCGCTGCTGCAAGCCCTGCGGCGGTGTTCTGACTGAG 846
Db      260 TPrpArgAspTyrIleAsnGlySerLeuLeuGluAlaLeuLysGlyValPheIleThrAsp 279
Qy      847 GCCCTGCGAGAGGCTGTGGCGCGGAGGAGCTTGTGCTGCTGTCAGTGGATGAGAGCT 906
Db      280 SerLeuLysGlnAlaValGlyHisGlnAlaIleLysLeuLeuValAsnValAspGlu 299
Qy      907 GAC 909
Db      300 Asp 300

RESULT 13
US-10-001-254-8
; Sequence 8, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Stehner-Liewen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-254-8

Alignment Scores:
Pred. No.: 2,36-24 Length: 101
Score: 514.00 Matches: 100
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 1
Query Match: 14.95% Indels: 0
DB: 14 Gaps: 0

US-10-030-271-1 (1-909) x US-10-001-254-8 (1-101)
```


QY 172 -----GGCTTAGCCCGGCGCCGCGCAGCGGC 195
Db 80 ArgMetGlnProGlnIuysGluGlyCyaserIleAlaSerArgHisArgProIuysSerGly 99
QY 196 CTAGAGCTCCTGCTGAGAGCTGAGCGCCGCGGAGAGCAACTGCGGCTG 255
Db 100 ValGluLeuLeuGluLeuGluArgGlyTyrLeuSerAspAlaAsnLeuArgPro 119
QY 256 CTGGGCACTCTGCGCGTGTGCGCCGCGCAGCACTGCGCAGCTGCGCGCAG 315
Db 120 LeuLeuGlnLeuLeuArgIleLeuThrArgHisAspValLeuProPheValSerGlnIuys 139
QY 316 CGGCGCCGCGCCAGTGTCTCA 336
Db 140 LysArgArgThrValSerPro 146

Search completed: March 22, 2005, 02:47:20
Job time : 114.904 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 21, 2005, 23:44:25 ; Search time 126.485 Seconds
(without alignments)
5558.996 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 1716
Sequence: 1 atggcgcctaccggcgtcgac.....tcagtcgtcagtgagcgtcgac 909

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPRO_pool_p/US10030271/runat_21032005_153457_26448/app_query.fasta_1.3150
-DB=Geneseq_16Dec04 -QMT=fastcan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030271 @CGN 1.1 445 @runat 21032005 153457 26448 -NCPU=6 -ICPU=3
-NO_MMAR -LARGOQUERY -NEG_SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Geneseq_16Dec04:*
2: Geneseqp1980s:*
3: Geneseqp1990s:*
4: Geneseqp2000s:*
5: Geneseqp2001s:*
6: Geneseqp2002s:*
7: Geneseqp2003as:*
8: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1550	90.3	303	4	AAB60386 Human apo
2	1550	90.3	326	4	AAB60387 Human apo
3	1543	89.9	318	5	AAE24860 Human DED
4	1543	89.9	318	5	AAE38903 Human DED
5	1543	89.9	326	5	ABBO7263 Human apo
6	1543	89.9	326	7	ADC79260 Human DED
7	1543	89.9	366	4	AAE73852 Human COL
8	1459	85.0	361	5	ABBO6038 Human NS
9	1459	85.0	368	5	ABBO6039 Human NS
10	1430	83.3	319	3	AAE58420 Lung can

11	1413	82.3	304	3	AAB15551
12	987	57.5	242	4	AAE25705
13	955	55.7	277	4	AAE39805
14	935.5	54.5	217	4	AAE41591
15	622.5	36.3	318	4	AAE94040
16	614.5	35.8	318	2	AAE90108
17	614.5	35.8	318	3	AAE51022
18	614.5	35.8	318	5	AAE26087
19	613.5	35.8	318	4	AAE93016
20	608.5	35.5	318	2	AAE90109
21	608.5	35.5	318	3	AAE51023
22	608.5	35.5	318	5	AAE26088
23	514	30.0	101	5	AAE24855
24	514	30.0	101	5	AAE38898
25	230.5	14.1	19938	6	ABP76680
26	221.5	13.5	369	7	ABO76648
27	217	13.2	1034	7	ABO80175
28	212	12.4	19938	6	ABE98398
29	210.5	12.9	309	7	ABO76640
30	210.5	12.9	19938	6	ABP76680
31	210	12.8	273	7	ABO74905
32	210	12.8	561	2	AAE37739
33	210	12.8	561	2	AAE93249
34	210	12.8	561	2	AAE57650
35	210	12.8	1040	4	ABG14734
36	209	12.8	1706	7	ABO83014
37	205.5	12.0	684	4	ABG14126
38	205	12.5	1411	7	ABO79662
39	204.5	12.5	1051	7	ADC31578
40	204	11.9	501	7	ABO74053
41	204	11.9	630	7	ABO76880
42	204	12.5	869	7	ABO73544
43	201.5	11.7	957	7	ABO71662
44	201	11.7	456	7	ABO70671
45	201	12.3	777	2	AAE37740

ALIGNMENTS

RESULT 1	
ID AAB60386	standard; protein; 303 AA.
AC AAB60386;	
DT 24-APR-2001	(first entry)
DE Human apoptosis-associated factor NT2RM1000558 (portion), SEQ ID NO:2.	
KW Human; apoptosis-associated factor; NT2RM1000558; death effector domain;	
KW DED; caspase family cleavage domain; pro-apoptotic; drug screening;	
KW cell proliferation; ischaemic disease; chronic viral disease.	
OS Homo sapiens.	
PN WO200104300-A1.	
PD 18-JAN-2001.	
PF 06-JUL-2000; 2000WO-JP004516.	
PR 08-JUL-1999; 99JP-00194179.	
PR 18-OCT-1999; 99US-0159566P.	
PA (HELI-) HELIX RES INST.	
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;	
DR WPI, 2001-138348/14.	
DR N-PSDB; AAF27407.	
XX Polynucleotide encoding an apoptosis-associated factor protein with death	
PT effector domain and caspase family-cleavage domain, useful in regulating	

PT diseases with cell proliferation.

XX Claim 1, Page 43-44; 53pp; Japanese.

XX The invention relates to a novel human apoptosis-associated factor (AAB60386, AAB60387), designated NT2RM1000558, which contains a death effector domain (DED) and a caspase family cleavage domain and is capable of inducing apoptosis in cells. The invention also relates to nucleic acids encoding the protein (AAB27407, AAB27408); variants of the protein (particularly dominant negative variants); vectors and host cells comprising a nucleic acid which encodes an apoptosis-associated factor of the invention; the recombinant production of the protein; an antibody against the protein; and methods of screening for compounds which can regulate apoptosis. The apoptosis-related factor is useful in regulating diseases associated with cell proliferation and in screening drug candidates e.g., for regulating cell proliferation or cell death in ischaemic diseases and chronic viral diseases. The present sequence CC represents a substantial proportion of the human apoptosis-associated CC factor NT2RM1000558

XX Sequence 303 AA;

Alignment Scores:

Pred. No.:	1.05e-109	Length:	303
Score:	1550.00	Matches:	303
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	90.33%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-1 (1-909) x AAB60386 (1-303)

QY 1 ATGGCGGTATCCGGGTGACCCCGCCGCTGCTGGAGAGATGATGCTTGACTAC 60
 Db 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluAspGluCysLeuAspTyr 20
 QY 61 TAGCGGATGCTGCTGCTTCAACCGTATTTGAGAGTGGTGGCGGCAACGACGATGC 120
 Db 21 TyrGlyMetLeuSerLeuHisArgMetPheGluValAlaGlyGlyGlnLeuThrGluCys 40
 QY 121 GAGCTGAGAGCTCCGGGCTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTGACC 180
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60
 QY 181 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGAGCATGGCGGAG 240
 Db 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluAlaArgArgGlyGlnCysGlyGlu 80
 QY 241 AGCAACTGCGGCTGCTGGGCAACTCTGCGCGCTGCTGGCCCGCCACGACTGTGCCG 300
 Db 81 SerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 301 CACCTGGCGGCAAGCGCGCGCGCGCGAGTGTCTCAGAACGCTATGACTTGACCTCC 360
 Db 101 HisLeuAlaArgGlySerArgArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
 QY 361 AGCTCTTCAAGAGAGAGAGAGAGGAGTGGCTGCTGCGCGCTGCGGAGTCAAGCATTTCTGCA 420
 Db 121 SerSerSerTyrSerArgThrGluGlySerCysArgArgArgArgGlnSerSerSerSerAla 140
 QY 421 AATTCTCAGCAGGGGTCACTGGAGAGACAGCTCCCTCCCAACCAAGCGGCGAGGAGT 480
 Db 141 AsnSerGlnGlnGlnGlnTyrGlnThrGlySerProProThrTyrSerArgAlaArgSer 160
 QY 481 CGGGGCGCGCGCGAGTGGTGGTGGCGAGCGGCGGAGAGGCGCGCGAGCGGCGAG 540
 Db 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 541 CAGCAGTCAGAGCGCGCGCAGACTTCTCTGAAGGCAAGTGAACCTTGACATCCGCTC 600
 Db 181 GlnGlnSerGluProAlaArgProSerSerSerGluGlyValThrCysAspIleArgLeu 200
 QY 601 CGGGTTGAGACAGAGTACTGGAGCATGGCCAGCCTTGAGAGCGGCGCTGGCATCCCGG 660

Db 201 ArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGluGlnGlyValAlaSerArg 220
 QY 661 CGGCCCCAGGCGCTGGCGCGCGAGCTGACGCTGTTTGGCGAGGCCACCCAGTGTCCGC 720
 Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 721 TCAGAGGACCTGGGCTCTGTGTTGTGACATCAATTCAGAGTCTCCATCTGAGAC 780
 Db 241 SerArgAspLeuGlySerValValCysAspIleLysPheSerGluLeuSerTyrLeuAsp 260
 QY 781 GCCTTGTGGGGCAGACTACCTGAGTGGCGCGCTGCTCAGAGCCCTGCGGGCGCTTCTTG 840
 Db 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 841 ACTGAGCCCTGGAGAGGCTGTGGGCGGAGGCTGTTGCGCTGCTGTCAGTGGAT 900
 Db 281 ThrGluAlaLeuArgGluAlaValAlaGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 901 GAGGCTGAC 909
 Db 301 GluAlaAsp 303

RESULT 2

ID AAB60387 standard; protein; 326 AA.

AC AAB60387;

DT 24-APR-2001 (first entry)

DE Human apoptosis-associated factor NT2RM1000558, SEQ ID-NO.4.

KW Human; apoptosis-associated factor; NT2RM1000558; death effector domain;

KW DED; caspase family cleavage domain; pro-apoptotic; drug screening;

KW cell proliferation; ischaemic disease; chronic viral disease.

OS Homo sapiens.

PN WO200104300-A1.

PD 18-JAN-2001.

PF 06-JUL-2000; 2000MO-JD004516.

PR 08-JUL-1999; 99JP-00194179.

PR 18-OCT-1999; 99US-0159586P.

PA (HELI-) HELIX RES INST.

PI Oka T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;

DR WPI; 2001-138348/14.

DR N-FSDB; AAB27408.

PT Polynucleotide encoding an apoptosis-associated factor protein with death

PT effector domain and caspase family-cleavage domain, useful in regulating

PT diseases with cell proliferation.

PS Claim 3; Page 47-48; 53pp; Japanese.

XX The invention relates to a novel human apoptosis-associated factor

XX (AAB60386, AAB60387), designated NT2RM1000558, which contains a death

XX effector domain (DED) and a caspase family cleavage domain and is capable

XX of inducing apoptosis in cells. The invention also relates to nucleic

XX acids encoding the protein (AAB27407, AAB27408); variants of the protein

XX comprising a nucleic acid which encodes an apoptosis-associated factor

CC ischaemic diseases and chronic viral diseases. The present sequence
 CC represents the human apoptosis-associated factor NT2RM1000558
 XX
 SQ Sequence 326 AA;

Alignment Scores:
 Pred. No.: 1.06e-109 Length: 326
 Score: 1550.00 Matches: 303
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.33% Indels: 0
 DB: 4 Gaps: 0

US-10-030-271-1 (1-909) x AAB60387 (1-326)

QY 1 ATGGCGTATCCGGGTCGACCCCGGCGCTGCGAGAGATGAGTGCCTGACTAC 60
 Db 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGlnGluAspGluCysLeuAspTyr 20
 QY 61 TAGCGGATGCTGTGCTTACCGATGTTTTCGAGGTGTGGCGGGCAACTGACCGATGC 120
 Db 21 TyrGlyMetLeuSerLeuHisArgGlyPheGluValValGlyGlyGlnLeuThrGluCys 40
 QY 121 GAGCTGAGCTCTCGGCTTTCTGCTGATGAGGCTCCTGGCGCGCCGCGAGGCTTACGC 180
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60
 QY 181 CGGGCCCGCAGCGGCGCTAGAGCTCCTGCTGAGCTGAGCGCGCGGCGGAGTGGCGGAG 240
 Db 61 ArgAlaAspSerGlyLeuGluLeuLeuLeuGluLeuGluAlaArgArgGlyGlnCysGlyGln 80
 QY 241 AGCAACCTGCGGCTGCTGCGGCAACTCCTGCGCGCTGCTGCGCGCGCACGACTGTGCGG 300
 Db 81 SerAsnLeuAlaGlyLeuGlnGlyGlnLeuLeuAlaArgHisAspLeuLeuPro 100
 QY 301 CACCTGGCGCGCAGCGGCGCGCGCATGCTCTCCAGAACCTATAGCTATAGGACCTTCC 360
 Db 101 HisLeuAlaArgGlyAspArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
 QY 361 AGCTCTTCAAGAGAGACAGAGGATGAGTGCCTGCGCGCTGCGGCGAGTCAAGAGATTCGCA 420
 Db 121 SerSerSerIysAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140
 QY 421 AATTCTCAGCAGGGTCACTGTGAGAGACAGGCTCCCCCAACCAAGCGGCGGCGGAGT 480
 Db 141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 QY 481 CGGGCGCGCGCGCATGCTGTGCTGCGAGAGCGCGGAGAGGGCGCCCGCGGACCCGAG 540
 Db 161 ArgGlyIysArgProSerGlyGlyAlaAspArgArgArgGlyAlaProAlaAlaProGln 180
 QY 541 CAGAGTCAAGAGCGCGCGGAGCTTCTCTGAGAGCAAGGAGAGCTGTGAGATCCGGGCTC 600
 Db 181 GlnLeuSerGluProAlaAspArgProSerSerGlnGlyGlyValThrCysAspIleArgLeu 200
 QY 601 CGGGTTCAGCAGAGTACTGCGAGCATGCGGCGGCTTTCGAGCAGGCGGCTGACATCCCG 660
 Db 201 ArgValAlaArgAlaGluTyrCysGlnHisGlyProAlaLeuGlnGlnGlnGlnGlnGln 220
 QY 661 CGGGCGCGCGCGCTGCGCGCGGAGTGAAGCTTTGGGCAAGCGGCAAGCTGTGCGG 720
 Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 721 TCAAGGAGCCTGGGCTGTGCTGTTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGC 780
 Db 241 SerArgAspLeuGlySerValValCysAspIleIysPheSerGlnLeuSerTyrLeuAsp 260
 QY 781 GCCTTCTGGGGCGACTACTAGTGGCGCGCTGCTGAGCGCGCTGCGGGCGGTGTTCTCG 840
 Db 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 841 ACTGAGGCCCTGCGAGAGGCTGTGGCGCGGAGGCTGTGCTGCTGCTGCTGCTGCTGAT 900

Db 281 ThrGluAlaLeuAspGlnAlaValGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 901 GAGGCTGAC 909
 Db 301 GluAlaAsp 303

RESULT 3
 AAE24860
 ID AAE24860 standard; protein; 318 AA.
 XX
 AC AAE24860;
 XX
 DT 22-OCT-2002 (first entry)
 XX
 DE Human DED4 (death effector domain) protein.
 XX
 KW Human; death domain; DD; death effector domain; DED; Chlamydia infection; NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis; inflammation; allergy; autoimmunity; allograft rejection; cell division; immune-based pathology; fibrosis; arthritis; graft versus host disease; immunosuppressive; gene therapy; antisense therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 53..74 /note="Nuclear localisation sequence"
 FT
 PN WO200240680-A2.
 XX
 PD 23-MAY-2002.
 XX
 PE 15-NOV-2001; 2001WO-US044844.
 XX
 PF 17-NOV-2000; 2000US-00715893.
 PR 29-JUN-2001; 2001US-0301889P.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Read JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Steiner-Liwen F;
 XX
 DR WP1; 2002-500222/53.
 DR N-PSDB; AAD40080.
 XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
 XX
 PS Claim 3; Page 186-187; 209pp; English.

The invention relates to an isolated polypeptide comprising a death domain (DD), death effector domain (DED) or NB-ARC domain. The invention is useful for identifying a binding agent, preferably a protein or a drug that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC domain from DAP3, IRAK4, CTDP (Chlamydia trachomatis DD protein), DED4 or NID (NGFR-interacting Death Domain), with a candidate binding agent and detecting the association of the domain and the candidate binding agent, by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or chemical crosslinking, nuclear magnetic resonance (NMR), mass spectroscopy (MS) and FPA. The invention is useful for modulating the level of a cell process such as cell proliferation, cell adhesion, cell stress responses, responses to microbial infection and B cell immunoglobulin class switching, in particular apoptosis within a cell. Antibody specifically reactive with CTDP DD of C. trachomatis, C. muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the CTDP DD protein is useful for detecting a Chlamydia infection. The invention is useful for modulating the activity of oncogenic proteins, for treating a pathology caused by the oncogenic proteins and for treating bacterial infections by modulating the activity of bacterial proteins. The protein and antibody specific for it are useful for discovery of drugs that suppress infection, inflammation, allergy, sepsis, autoimmunity, allograft rejection and other diseases. The protein

CC is useful for treating immune-based pathologies, pathologies associated with cell division, inflammatory diseases such as sepsis, fibrosis, arthritis, graft versus host disease. The invention is used in antisense therapy and gene therapy. The present sequence is human DED4 protein

XX Sequence 318 AA:

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score: 1543.00	318	302	0	1	0	0
Percent Similarity: 99.67%						
Best Local Similarity: 99.67%						
Query Match: 89.92%						

US-10-030-271-1 (1-909) x AAE24860 (1-318)

```

QY 1 ATGGCGCTATCCGGGATGACCCCGGCGGAGGAGATGATGCTGAGTAC 60
DB 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluAspGlnCysLeuAspTyr 20
QY 61 TACGGATGCTGTCGCTTCAACCGTATGTCGAGGTGAGCGGCAACTGACCGAGTGC 120
DB 21 TyrGlyMetLeuSerLeuHisArgMetPheGluValGlyGlnLeuThrGlnCys 40
QY 121 GAGCTGAGGCTCTCGGCTTTTCTGCTGATGAGGCTCTGAGCGCGCGGAGCTTAGCC 180
DB 41 GluLeuGlnLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaIleGlyLeuAla 60
QY 181 CGGGCCCGGAGCGGCGCTAGAGCTCTGAGGAGGAGGCGGCGGAGTGGCGGAG 240
DB 61 ArgAlaArgSerGlyLeuGlnLeuLeuGlnLeuGlnLeuArgValLeuAlaArgHisAspLeuPro 80
QY 241 AGCAACCTGCGGCTGCTGAGGAGCAACTCTGCGCTGCGCGCGCGCACTGCTCCG 300
DB 81 SerAsnLeuArgLeuLeuGlnLeuLeuLeuArgValLeuAlaArgHisAspLeuPro 100
QY 301 CACCTGGCGCGCAAGCGGCGCGCGCGAGTCTTCAGAGCGCTATGACACTTCC 360
DB 101 HisLeuAlaArgLysArgArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
QY 361 AGCTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 121 SerSerSerLysArgLysArgLysArgLysArgLysArgLysArgLysArgLys 140
QY 421 AATTCTCAGAGGAGTCTGAGTGGAGACAGGCTCCCGCCCAACCAAGCGGCGGAGT 480
DB 141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
QY 481 CGGGCGCGGCGCGAGTGGTGGCCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
QY 541 CAGCAGTCAGAGCGCGCGAGACTCTCTTGAAGGCAAGTGAAGCTTGAATCCGGCTC 600
DB 181 GlnGlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGlnGln 200
QY 601 CGGCTTCAGACAGAGTACTGAGCATGAGGCGGAGCTTGAAGGAGGAGGAGGAGGAG 660
DB 201 ArgValAlaArgAlaGlnTyrCysGlnHisGlyProAlaLeuGlnGlnGlnGlnGln 220
QY 661 CGGCGCGGCGGCTGAGCGGCGGAGCTGAGCGTGTGTTGGGCGGAGCGGAGTCTGCG 720
DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
QY 721 TCAGAGGAGCTGGGCTCTGTGTTTGTGACATCAAGTTTCAGAGCTCTCTATCTGAC 780
DB 241 SerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrLeuAsp 260
QY 781 GCGCTTCGGGCGGAGCTGAGTGGGCGGCGGCTGTCGAGGCGGCGGCGGCGGCTTCTG 840
DB 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280

```

QY 841 ACTGAGCCCTGCGAGAGGCTGNGGCGGAGGCTGTTGCGCTGCTGAGTGGAT 900
 DB 281 ThrGluAlaLeuArgGlnAlaValGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 901 GAGCTGAC 909
 DB 301 GluAlaAsp 303

RESULT 4

AAE38903

ID AAE38903 standard; protein; 318 AA.

XX AAE38903;

AC 18-DEC-2003 (first entry)

DE Human DED4 protein.

KW Human, death Domain; DD; death effector domain; DED; cell proliferation;

KW Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;

KW neural growth factor receptor-interacting death domain; cell adhesion;

KW vasotropic; microbial infection; inflammation; allograft rejection; CTDD;

KW cell stress response; benign prostatic hyperplasia; antibacterial; NIDD;

KW apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;

KW neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;

KW Keloid.

XX Homo sapiens.

OS

FT Key Location/Qualifiers

FT Domain 12..112 /note="Death effector domain of DED-4"

FT Peptide 53..74 /note="Nuclear localisation sequence"

XX US2003049702-A1.

XX 13-MAR-2003.

XX 15-NOV-2001; 2001US-00001254.

XX 17-NOV-2000; 2000US-00715893.

XX 17-NOV-2000; 2000US-0367360P.

XX 29-JUN-2001; 2001US-0301889P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (PAWL/) PAWLOWSKI K.

XX (FIOR/) FIORENTINO L.

XX (LEES/) LEE S H.

XX (ROTH/) ROTH W.

XX (STEN/) STENNER-LIEWEN F.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC that modulates the association of a DD, DED or NB-ARC domain with protein
 CC that binds the DD, DED or NB-ARC domain. The invention is also useful for
 CC modulating the level of cell process such as apoptosis, cell adhesion,
 CC cell proliferation, cell stress responses, responses to microbial
 CC infection and B cell immunoglobulin class switching. DDE, DEDs and NB-ARC
 CC domains and/or anti-DD, anti-DDD or anti-NB-ARC domain antibodies are
 CC useful for discovery of drugs that suppress infection, autoimmunity,
 CC inflammation, allergy, allograft rejection, sepsis and other diseases.
 CC DD or NB-ARC domain proteins are used to treat infection, allergy,
 CC autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
 CC hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
 CC inflammatory hyperplasia and smooth muscle cell proliferation in arteries
 CC following balloon angioplasty (restenosis). The invention is also used in
 CC antibody therapy and gene therapy. The present sequence is human DED4
 CC protein. The DED4 gene is located on chromosome 19
 XX

XX Sequence 318 AA:

Alignment Scores:

Pred. No.:	3,66-109	Length:	318
Score:	1543.00	Matches:	302
Percent Similarity:	99.67%	Conservative:	0
Best Local Similarity:	99.67%	Mismatches:	1
Query Match:	89.92%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-1 (1-909) x AAB38903 (1-318)

QY 1 ATGGCGCTATCCGGGTGCAACCCCGGCTGCTGGAGAGATGAGTCTTGACCTAC 60
 DB 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluGluAspGluCysLeuAspTyr 20
 QY 61 TAGGGAGTGTGCTGCTCACCGTATGTTCCGAGGGTGGGGGGCACTGACGAGTGC 120
 DB 21 TyrluYeuLseuSerLeuHisArgIuePheGluValAlaGlyGlnuethrGluCys 40
 QY 121 GAGCTGAGAGCTCTGCTTCTGCTGTGATGAGGCTCTCTGCGCGCCGAGGCTTACC 180
 DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlnuethr 60
 QY 181 CGGGCCCGGCGGCTTACAGCTCTGCTGAGCTGAGGCGCGGCGGAGTGGCGGAG 240
 DB 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluArgGlyGlnCysAspGlu 80
 QY 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGCGCTGCGCGCCGAGGAGCTGTGCG 300
 DB 81 SerAsnLeuArgLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 301 CACCTGGCGCGCAAGCGCGCGGCACTGTCTCCAGAACGCTATAGTACCTACCTCC 360
 DB 101 HisLeuAlaArgLysArgArgProAlaSerProGluAlaGlyTyrSerTyrGlyThrSer 120
 QY 361 AGCTCTTCAAGAGAGAGAGAGGTAAGTCCGCTGCGCGCTGCGCAAGAGTTCTGCA 420
 DB 121 SerSerSerLysArgThrGluGlySerCysArgArgArgGlnSerSerSerSerAla 140
 QY 421 AATCTCAGCAGGCTGATGAGAGACAGGCTCCCGCCCAACCAAGCGGACGCGGAGT 480
 DB 141 AsnSerGlnGlnGlnGlnTyrGlnTyrGlnTyrGlnTyrGlnTyrGlnTyrGlnTyr 160
 QY 481 CGGGCCCGGCGGAGTGTGTGTCAGAGCGGCGCGAGAGGGGCGCCAGCGCCAGCCAG 540
 DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 541 CAGCAGTCAAGCGCGCGAGACTTCTCTGAGAGCAAGTGAAGCTGTGAGATCCGAGTC 600
 DB 181 GlnGlnSerGluProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeu 200
 QY 601 CGGCTTCAAGAGAGTACTGCGAGATAGGGCCAGGCTTGGAGCAGGCGGTGACATCCGG 660
 DB 201 ArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGluGlnGlnValAlaSerArg 220
 QY 661 CGGGCCCGGCGGCTGCGCGGAGCTGACGTTGTTGGGAGGCGCCAGCGAGTGTGCGC 720

DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValaPheGlyGlnAlaThrAlaValLeuArg 240
 QY 721 TCAAGGAGACCTGGGCTGCTGTTGTTGACATCAAGTTCAGAGCTCTCTATCTGAGC 780
 DB 241 SerArgAspLeuGlySerValaCysAspIleLysPheSerGluLeuSerTyrLeuAsp 260
 QY 781 GCCTTCTGGGCGCATCACTGAGTGGCGCCCTGCTGAGGCGCTGCGGGCGTTCCTG 840
 DB 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValaPheLeu 280
 QY 841 ACTAGGCGCTTGGCAGAGGCGCTGTGGCGCGGAGGCTGTGCTGCTGCTGATGAT 900
 DB 281 ThrGluAlaLeuArgGluAlaValaGlyArgGluAlaValaArgLeuLeuValSerValaAsp 300
 QY 901 GAGGCTGAC 909
 DB 301 GluAlaAsp 303

RESULT 5

ABB07263 ABB07263 standard; protein; 326 AA.

AC ABB07263;

DT 26-MAR-2002 (first entry)

XX Human apoptosis regulator (APRG) polypeptide (Incyte ID. 3102521C01).

KW APRG; apoptosis regulator; cytosolic; antiatherosclerotic; osteopathic;

KW antiatherosclerotic; hepatocytic; antiproliferative; antihemantic; human;

KW antiallergic; antiasthmatic; antiasthmatic; antihypertensive; cancer;

KW antiinflammatory; antidiabetic; antipain; nephrotoxic; ophthalmological;

KW immunosuppressive; dermatological; antitumor; antihemantic; fungicide;

KW antitumor; antibacterial; virucide; antiparasitic; protozoicide;

KW triamizolizer; valnerary; gynecological; vasoactive; gene therapy.

OS Homo sapiens.

XX WO200192527-A2.

PN 06-DEC-2001.

PF 30-MAY-2001; 2001WO-US017581.

PR 01-JUN-2000; 2000US-0209407P.

PR 30-NOV-2000; 2000US-0250326P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Azimzai Y, Yue H, Burford N, Ding L, Elliott VS;

PI Patterson C, Baughn MR;

XX WPI; 2002-114350/15.

DR N-PSDB; ABA94362.

XX Novel human apoptosis regulator polypeptides and polynucleotides for

PT diagnosing, preventing, and for identifying modulators of therapeutic use.

PT reproductive disorders and for identifying modulators of therapeutic use.

XX Claim 1; Page 97-98; 103pp; English.

XX The invention provides human apoptosis regulator (APRG) polypeptides and

CC polynucleotides. The APRG polypeptides, polynucleotides and modulators

CC are useful for diagnosis, treatment and prevention of cell proliferative,

CC glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
 CC erythema multiforme, systemic sclerosis, ulcerative colitis, haemodialysis,
 CC uveitis; viral, bacterial, fungal, parasitic, protozoal, helminthic
 CC infections and trauma. Reproductive disorders include disorders of
 CC prolactin production, infertility, endometriosis, polycystic ovary
 CC syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,
 CC disruptions of spermatogenesis, cancer of testis and prostate, impotence,
 CC carcinoma of male breast and gynecomastia. The APRG polynucleotides are
 CC useful for creating knockin humanized animals or transgenic animals to
 CC model human disease and to detect and quantify gene expression in
 CC biopsied tissues in which expression of APRG is correlated with disease.
 CC APRG fragments of it and antibodies specific for APRG are useful as
 CC elements on a microarray which is useful to monitor or measure protein-
 CC protein interactions, drug-target interactions and gene expression
 CC profiles. The present sequence represents a human APRG polypeptide
 CC
 XX

SO Sequence 326 AA;

Alignment Scores:

Pred. No.:	3,62e-109	Length:	326
Score:	1543.00	Matches:	302
Percent Similarity:	99.67%	Conservative:	0
Best Local Similarity:	99.67%	Mismatches:	1
Query Match:	89.92%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-1 (1-909) x ABB07263 (1-326)

QY 1 ATGGCGGATCCGGAGTACCGCGCCCGTGGAGGAGATGAGTACCTGAGTAC 60
 DB 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluIleuAspGluCysLeuAspTyr 20
 QY 61 TACGGAGTGGTGGCTTCAACCGTATGTTGAGAGTGGGCGGCGCACTGACCGAGTGC 120
 DB 21 TyrGlyMetLeuSerLeuHisArgMetPheGluValValGlyGlnLeuThrGluCys 40
 QY 121 GAGCTGAGAGTCTCTGGCTTCTGCTGGATGAGGCTCTCGCGCCGCGGAGGCTTACCC 180
 DB 41 GluLeuGluIleuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60
 QY 181 CGGGCCCGGCGCGCTAGAGGCTCTGCTGAGAGTGGAGCGCGGCGGAGTGGCGGAG 240
 DB 61 ArgAlaArgSerGlyLeuGluIleuLeuGluIleuGluIleuArgGlyGlnCysAspGlu 80
 QY 241 AGCAACCTGGCGCTGCTGGGCGCACTCTCGCGTGTGGCCCGCCACGACTGTGCGC 300
 DB 81 SerAsnLeuArgLeuLeuGluGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 301 CACCTGGCGCGCAAGCGCGCCCGCGCAGTGTCTTCAGAAAGCTTATGGCACTCC 360
 DB 101 HisLeuAlaArgIysArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
 QY 361 AGCTCTCAAGAGAGAGAGGAGTGGCGTGGCGCGGAGTGGAGAGTGGTTCGA 420
 DB 121 SerSerSerIysArgThrGluGlySerCysArgArgArgArgGlnSerSerSerAla 140
 QY 421 AATTTCAGAGAGGTGAGTGGAGAGAGGCTCCCGCCCAACGAGCGGAGCGGAGT 480
 DB 141 AsnSerGlnGlnGlnTrpGluThrGlySerProThrIysArgGlnArgArgSer 160
 QY 481 CGGGGCGCGCGCGAGTGGTGGTCCAGACGCGCGGAGAGAGGCGCGCCAGCCCCAG 540
 DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 541 CAGAGTCAAGAGCGCGCGCAAGCTTCTGAGAGGCAAGGAGCTGTGACATCCGGCTC 600
 DB 181 GlnGlnSerGluProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeu 200
 QY 601 CGGGTTTCAGACAGATGATCTGAGCATGGCGCAGCTTGGAGCAGGCGTGGCATCCCGG 660
 DB 201 ArgValArgIleArgIleGluTyrCysGlnHisGlyProAlaLeuGluGlnGlnValAlaSerArg 220
 QY 661 CGGCGCCAGGCGCTGGCGCGAGAGCTGAGACGTGTTGGGAGGCGCACCGAGTGGCGC 720

DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 721 TCAGAGGACCTGGGCTCTGTGTTTGGATCATCAAGTCTCAGAGCTCTCTTACGAGC 780
 DB 241 SerArgAspLeuGlySerValValValCysAspIleIlePheSerGlnLeuSerTyrLeuAsp 260
 QY 781 GCCTTTCGGGCGACATCACTGAGTGGCGCGCTCTGCGAGGCGCTGGCGGCGTTCCTG 840
 DB 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 841 ACTGAGGCGCTGGAGAGGCTGGGCGCGGAGGCTGTTCGCTGCTGGTCACTGGAT 900
 DB 281 ThrGlnAlaLeuArgGluAlaValGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 901 GAGCTGAC 909
 DB 301 GluAlaAsp 303

RESULT 6

ADCT9260
 ID ADCT9260 standard; protein; 326 AA.

AC ADCT9260;

AC 01-JAN-2004 (first entry)

DE Human DEDD2 protein SEQ ID NO:2.

XX human, death effector domains containing DNA-binding protein;

KW DED2-containing DNA-binding protein; DEDD2, cell death; gene therapy;

KW cytostatic; cancer; chronic myeloid leukaemia.

OS Homo sapiens.

PN WC02003054195-A1.

PD 03-JUL-2003.

PF 20-DEC-2002; 2002MO-JP013371.

XX 20-DEC-2001; 2001JP-00387854.

PR 18-JUL-2002; 2002JP-00209458.

XX (MORG) MORINAGA MILK IND CO LTD.

PA Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;

XX WPI, 2003-569246/53.

DR N-PSDB; ADCT9259.

XX DNA encoding cell death proteins for treatment of kidney, large intestine

PT and prostate cancers and leukemia.

XX Claim 1, Page 20-21; 26pp; Japanese.

XX The present sequence represents a human death effector domains (DED)

CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that

CC causes cell death. Also described: (1) primer and probe for investigation

CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has

CC cytosolic activity. DEDD2 can be used in the diagnosis and treatment of

CC cancers of the kidney, large intestine and prostate, and acute and

XX chronic myeloid leukaemia.

SO Sequence 326 AA;

Alignment Scores:

Pred. No.:	3.62e-109	Length:	326
Score:	1543.00	Matches:	302
Percent Similarity:	99.67%	Conservative:	0
Best Local Similarity:	99.67%	Mismatches:	1
Query Match:	89.92%	Indels:	0
DB:	7	Gaps:	0

US-10-030-271-1 (1-909) x ADC79260 (1-326)

QY 1 ATGGCGCTATCCGGGTCGACCCCGGCGCTGCTGGAGAGATGAGTGCCTGAGTAC 60

DB 1 MetAlaLeuSerGlySerThrProAlaProCysTrpGluGluNabpGluCysAlaLeuAspTyr 20

QY 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCGGGCAAATGACCGAGTGC 120

DB 21 TyrGlyMetLeuSerLeuH1sArgMetPheGluValGlyGlyGlnLeuThrGluCys 40

QY 121 GAGCTGAGAGCTCCGGGCTTCTGCTGATGAGCTCTGAGAGCTGAGCGCGGAGGCTTGGCC 180

DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60

QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCTGAGCGCGGAGGCTGAGCGGAG 240

DB 61 ArgAlaArgSerGlyLeuGluLeuLeuGluGluGluGluGluGluGluGluGluGlu 80

QY 241 AGCAACCTGGCGGCTGCTGGGGCAAATCTCTGCGCGCTGCTGCGCGCACGACCTGCGCG 300

DB 81 SerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgH1sAspLeuLeuPro 100

QY 301 CACCTGGCGGCGCAGCGGCGCGCGCGCTGCTCTGAGAGCGCTTACCTATGCACTCC 360

DB 101 HisLeuAlaArgLysArgArgArgProValSerProGluArgLysTyrGlyThrSer 120

QY 361 AGCTCTCAAGAGAGAGAGAGAGAGAGTACGCTCGCGCTGCGCGCTGAGAGAGAGTTCGCA 420

DB 121 SerSerSerLysArgThrGluGlySerCysArgArgArgArgGlnSerSerSerAla 140

QY 421 AATTTCAGAGAGGTCAGTGGAGAGAGAGTCCCGCCCAACAGCGGAGCGGAGT 480

DB 141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160

QY 481 CGGGCG 540

DB 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgArgArgArgArgArgArg 180

QY 541 CAGAGCGAGAGCGCGCGCGCGCGCGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

DB 181 GlnGlnSerGluProAlaArgProSerSerGluGlyValThrCysAspIleArgLeu 200

QY 601 CGGAGTTCAGAGAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

DB 201 ArgValArgArgLysGluGlyCysGlnH1sGlyProAlaLeuGluGlnGlnGlnGlnGln 220

QY 661 CGGCG 720

DB 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240

QY 721 TCAAGGAGAGCTGGGCTGTGCTTGTGAGCATCAAGTTCAGAGCTCTCTATCTGAG 780

DB 241 SerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrLeuAsp 260

QY 781 GCCTTCCTGGGCGGAGTCTGAGTGGCGCGCGCTGCTGAGAGCGCTGCGCGGAGTTCCTG 840

DB 261 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuAlaGlyValPheLeu 280

QY 841 ACTGAGGCGCTGCGAGAGAGCTGTGGCGCGGAGGCTGTGCTGCTGCTGCTGCTGCTGAT 900

DB 281 ThrGluAlaLeuArgGluAlaValGlyArgGluAlaValAlaArgLeuLeuValSerValAsp 300

QY 901 GAGGCTGAC 909

DB 301 GluAlaAsp 303

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:4616.

DE Human: colon cancer; colon cancer antigen; diagnosis; detection;

KW colorectal carcinoma.

XX Homo sapiens.

OS WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US026524.

XX 29-SEP-1999; 99US-0157137P.

PR 03-NOV-1999; 99US-0163280P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Barash SC, Birse CE, Rosen CA;

PI WPI: 2001-235357/24.

DR N-PSDB; AAH33283.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers.

PS Claim 11; Page 6413-6414; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where the

CC proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene therapy

CC and vaccine production. N and P may be used in the prevention, diagnosis

CC and treatment of diseases associated with inappropriate P expression. For

CC example, N and P may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome

CC that affect the activity of P by expressing inactive proteins or to

CC supplement the patient's own production of P. Additionally, N may be used

CC to produce the colon cancer-associated Ps, by inserting the nucleic acids

CC into a host cell and culturing the cell to express the proteins. N and P

CC can be used in the prevention, diagnosis and treatment of colorectal

CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent

CC sequences used in the exemplification of the present invention. N.B.

CC Pages 666 to 682 and page 7053 of the sequence listing were missing at

CC time of publication, meaning no sequences are present for SEQ ID NO:1027

CC to 1052, 7921 and 7922

XX

XX Sequence 366 AA;

XX

XX Alignment Scores:

XX Pred. No.: 3,68e-109 Length: 366

XX Score: 1543.00 Matches: 302

XX Percent Similarity: 99.67% Conservative: 0

XX Best Local Similarity: 99.67% Mismatches: 1

XX Query Match: 89.92% Indels: 0

XX DB: 4 Gaps: 0

US-10-030-271-1 (1-909) x AAG73852 (1-366)

QY 1 ATGGCGCTATCCGGGTCGACCCCGGCGCTGCTGGAGAGATGAGTGCCTGAGTAC 60

DB 41 MetAlaLeuSerGlySerThrProAlaProCysTrpGluGluNabpGluCysAlaLeuAspTyr 60

QY 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCGGGCAAATGACCGAGTGC 120

DB 61 TyrGlyMetLeuSerLeuH1sArgMetPheGluValGlyGlyGlnLeuThrGluCys 80

QY 121 GAGCTGAGAGCTCCGGGCTTCTGCTGATGAGCTCTGAGAGCTGAGCGCGGAGGCTTGGCC 180

DB 81 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 100


```

QY 181 CGGGCCGCGACGGCCCTTGAAGCTCTGCTGAGAGCTGAGCCGCCGCGGACGTGGCGGAG 240
    |||||
DB 89 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluArgGluGlnCysAspGlu 108
    |||||
QY 241 AGCAACCTGGCGGCTGCTGGGCAACTCTGCGCTGCTGCGCCGCGACGACTCTGCGG 300
    |||||
DB 109 SerAsnLeuArgLeuGluGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 128
    |||||
QY 301 CACCTGGCGCGCAAGCGCGCGCCAGTGTCTCAGAAACCTATAGCTATAGGACCTCC 360
    |||||
DB 129 HisLeuAlaArgLys-AlaAlaGlyGlnCysLeuGlnHisAlaLeuAlaMetAlaProPr 148
    |||||
QY 361 AGCTTTCAAAGAGACAGAGAGGTAGTCCGCTGCGCGCTGGAGTCAAGCATGTTTCA 420
    |||||
DB 148 cAlaLeuGlnArgGlyGlnArgValAlaAlaValAlaValGlySerGlnAlaValLeuGlu 168
    |||||
QY 421 AATTCTCAGACGGGTCACTGGGAGACAGGCTCCCCCCCAACCAAGCGCGACGGCGGAGT 480
    |||||
DB 168 nileLeuSerArgPheSerGlyArgGlnAlaProGlnProSerGlyThr-AlaGluS 188
    |||||
QY 481 CGGGCCGCGCCAGTGTGTGTGTCAGACGCGCGCGAGAGGGGCGCCGACCCGAC 540
    |||||
DB 188 erGlyProAlaGlnTrpTrpCysGlnThrAlaAlaGluArg-AlaProAlaAlaProGln 207
    |||||
QY 541 CAGCAGTCAGAGCCCGCCAGACCTTCTTGAGGCAAAAGTGACCTGTGACATCCGCTC 600
    |||||
DB 208 GlnGlnSerGlnProAlaArgProSerSerGlyGlyValThrCysAspIleArgLeu 227
    |||||
QY 601 CGGGTTTGACAGAGTACTGAGACATGGGCGACCTTTGAGAGCGGGGCTGGCATCCCGG 660
    |||||
DB 228 ArgValArgAlaGlyLysGlyLysGlyLysGlyProAlaLeuGlnGlyValAlaSerArg 247
    |||||
QY 661 CGGGCCCGAGGCGCTGGCGCGCGACAGTGGATGTTTGGGCGAGCGACCGAGTGTGGGC 720
    |||||
DB 248 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 267
    |||||
QY 721 TCAAGGAGACCTGGCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGC 780
    |||||
DB 268 SerArgAspLeuGlySerValValCysAspIleLeuPheSerGlnLeuSerTyrLeuAsp 287
    |||||
QY 781 GCCTTTCGGGGCCACTACCTGAGTGGGCGCTGCTGCGAGCGCTCGGGGCGCTTCTCTG 840
    |||||
DB 288 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 307
    |||||
QY 841 ACTGAGGCGCTGGAGAGGCTGTGGGCGGGAGGCTGTGCTGCTGCTGAGTGGAT 900
    |||||
DB 308 ThrGlnAlaLeuArgGlnAlaValAlaGlyArgGln-LeuPheAlaCysTrpSerValTrpMe 327
    |||||
QY 901 GAGGCTG 907
    |||||
DB 327 CArgLeu 329
    |||||

RESULT 9
ABBO6039
ID ABB06039 standard; protein; 368 AA.
XX
XX ABB06039;
AC
XX
XX 10-MAY-2002 (first entry)
DT
XX
XX Human NS protein sequence SEQ ID NO:131.
DE
XX
XX Human; cytoeatic; osteopathic; gynaecological; neuroprotective;
KM antineumatic; antiarthritic; antiporiatic; ophthalmological; anti-HIV;
KM vasoectopic; antiarteriosclerotic; antiinflammatory; dermatological;
KM anorectic; muscular; antinfertilit; cardiovascular; anticoagulant;
KM antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;
KM anticoagulant; antidiabetic; tranquiliser; antidepressant; auroleptic;
KM gastrointestinal; vitnude; antituler; cerebroprotective; nootropic;
KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KM rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;

```

```

KM infertilit; cardiovascular disease; coagulation disease; hypertension;
KM ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KM gastric ulcer; Alzheimer's disease.
XX
XX Homo sapiens.
XX
XX MO200206315-A2.
XX
XX 24-JAN-2002.
XX
XX 17-JUL-2001; 2001MO-IL000653.
XX
XX 18-JUL-2000; 2000IL-00137345.
XX
XX 15-DEC-2000; 2000IL-00140354.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Mintz L, Freilich S, Bernstein J;
XX
XX WPI; 2002-155037/20.
XX
XX N-PSDB; ABL39693.
XX
XX One hundred and twenty eight novel nucleic acid sequences, useful for
XX treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
XX Claim 6; Page 151-152; 290pp; English.
XX
XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
XX encoding the proteins given in ABB06037 to ABB06164. The novel sequences
XX (NS) can have cytoeatic, osteopathic, gynaecological, neuroprotective,
XX antineumatic, antiarthritic, antiporiatic, ophthalmological, vitnude,
XX vasoectopic, antiarteriosclerotic, antiinflammatory, dermatological,
XX anorectic, muscular, anti-HIV, antinfertilit, cardiovascular, cardiant,
XX anticoagulant, antifibrinolytic, hypotension, antiasthmatic, tranquiliser,
XX immunomodulator, anticonvulsant, antidiabetic, cardiant,
XX antidiabetic, gastrointestinal, auroleptic, cerebroprotective,
XX nootropic and contraceptive activities. The NS can be used in vaccines,
XX gene therapy and antitense therapy. Nucleic acids, expression vectors and
XX antibodies from the present invention can be used for treating and
XX diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
XX diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
XX cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
XX glaucoma, obesity, muscular dystrophy, AIDS, infertilit, cardiovascular
XX disease, coagulation disease, ischaemia, hypertension, asthma, immune
XX disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
XX depression, schizophrenia, viral disease, gastric ulcers, stroke,
XX Alzheimer's disease and as a contraceptive
XX
XX
XX Sequence 368 AA;
SQ
Alignment Scores:
Pred. No.: 9,4e-103 Length: 368
Score: 1459.00 Matches: 293
Percent Similarity: 97.05% Conservative: 3
Best Local Similarity: 96.07% Mismatches: 5
Query Match: 85.02% Indels: 4
DB: Gaps: 0

US-10-030-271-1 (1-909) x ABB06039 (1-368)
QY 1 ATGGCGCTATCCGGGTGACCCCGCGCGCTGCTGGAGAGGATGAGTGCCTTGACTAC 60
    |||||
DB 36 MechAlaLeuSerGlySerThrProAlaProCysTrpGluGlnAspGluCysLeuAspTyr 55
    |||||
QY 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGCAACTGACCGAGTGC 120
    |||||
DB 56 TyGlyMetLeuSerLeuHisArgMetTyrGluValAlaGlyGlnLeuThrGluCys 75
    |||||
QY 121 GAGTGGAGGCTCTGCGCTTTCCTGCTGAGTGGAGGCTCTGCGCGCGGAGGCTTACGC 180
    |||||
DB 76 LysLeuGlnLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyTyrAla 95
    |||||

```

QY 181 CGGGCCCGCAGCGCCCTAGAGCTCTGCTGAGACTGAGAGCGCCGCGCAGTGGCGAG 240
DB 96 ArgAlaArgSerGlyLeuGlnLeuGlnLeuGlnLeuGlnArgGlnGlnCysAspGln 115
QY 241 AGCAACCTGCGGCTGCGGGGCAACTCTCTGCGGTGCGCCGACGACCTGCTGGCG 300
DB 116 SerAsnLeuArgLeuGlnGlnLeuLeuArgValLeuAlaArgHisAspLeuPro 135
QY 301 CACCTGCGCGCAGCGCGCCGCGCAGTGTCTCCAGAACCTATAGCTATGGCACCTCC 360
DB 136 HisLeuAlaArgLys-AlaAlaGlyGlnCysLeuGlnAsnAlaAlaAlaMetAlaProP 155
QY 361 AGCTCTTCAAAAGAGACAGAGGGTAGTCCCTGCGCGCTGGCAGTCAAGCACTTCCGA 420
DB 155 oAlaLeuGlnArgGlyGlnArgValAlaAlaValAlaGlySerGlnAlaValLeuGln 175
QY 421 AATTCTCAGAGGGTCAAGTGGAGAGACAGCTCCCGCCCAACAGAGCGCGCGAGT 480
DB 175 nileuSerArgPheSerGlyArgGlnAlaProProGlnProSerGlyThr-AlaGlnS 195
QY 481 CGGGCCCGCGCCAGTGGTGTGCGCAGCGCGCGAGAGAGGCGCCGACCGCAGCCGAG 540
DB 195 erGlyProAlaGlnTPTPTGCGlnThrAlaAlaGlnArg-AlaProAlaAlaProGln 214
QY 541 CAGCAGTCAAGCGCCGACACCTCTCTGAAAGCAAAGTACCTGTGACATCCGGCTC 600
DB 215 GlnGlnSerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeu 234
QY 601 CGGGTTCAGACAGATACCTGCGAGCATGGGCGAGCCTTGAGAGCGCGCGATCCCGG 660
DB 235 ArgValaArgAlaGlnTyrCysGlnHisGlyProAlaLeuGlnGlnGlyAlaAlaSerArg 254
QY 661 CGGCCCCGAGCGCTGGCGCGCGCAGCTGAGAGCTGTTGGGCGAGCCAGCAGTGTGCGC 720
DB 255 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 274
QY 721 TCAAGGAGCTGGGCTGTGTGTTTGACATCAAGTCTCAAGCTCTCTATCTGGAC 780
DB 275 SerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrLeuAsp 294
QY 781 GCCTTCGGGGGCGACTGAGTGGGCGCTGTGAGGCGCTGGCGGGCGTGTCTG 840
DB 295 AlaPheTyrPolaAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 314
QY 841 ACTGAGGCTGCGAGAGGCTGTGGCGCGAGAGGCTGTGCTGTGATCAAGTGTGAT 900
DB 315 ThrGlnAlaLeuArgGlnAlaValGlyArgGln-LeuPheAlaCysTyrSerValTyrPme 334
QY 901 GAGGCTG 907
DB 334 CArgLeu 336
RESULT 10
AAB58420
ID AAB58420 standard; protein; 319 AA.
XX
AC AAB58420;
XX
DT 14-MAR-2001 (first entry)
XX
DE Lung cancer associated polypeptide sequence SEQ ID 758.
XX
KW Human; lung cancer associated protein; neuroprotective; cyrostatic;
KW cardioactive; immunomodulatory; muscular active; vulnerary;
KW gastrointestinal; nephrotoxic; antiinfective; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.
XX
OS Homo sapiens.
XX
PN WO200055180-A2.
XX
PD 21-SEP-2000.

XX
PF 08-MAR-2000; 2000MO-US005918.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PA (ROSE-) ROSEN C A.
XX
PI Ruben SM;
XX
DR WPI; 2000-587514/55.
XX
DR N-PsDB; AAF18296.
XX
PT Lung cancer associated gene sequences, referred to as lung cancer
XX
PS Claim 11; Page 1275-1276; 1425pp; English.
XX
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cyrostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
CC general; nephrotoxic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 319 AA;
XX
Alignment Scores:
Pred. No.: 1,5e-100 Length: 319
Score: 1430.00 Matches: 278
Percent Similarity: 99.29% Conservative: 0
Best Local Similarity: 99.29% Mismatches: 1
Query Match: 83.33% Indels: 1
DB: 3 Gaps: 0
US-10-030-271-1 (1-909) x AAB58420 (1-319)
QY 1 ATGGCCCTATCCGGGTGACCCCGCGCCGCTGCGAGAGAGATAGTGCCTGACTAC 60
DB 41 MetAlaLeuSerGlySerThrProAlaProCysTyrGlnGlnAspIleCysLeuAspTyr 60
QY 61 TACGGATGCTGTCGCTTACACCGTATGTCGAGTGTGGGCGGCAACGACGAGTGC 120
DB 61 TyrGlyMetLeuSerLeuHisArgMetPheGlnValValGlyGlnLeuThrGlnCys 80
QY 121 GAGCTGAGCTCTGCGCTTCTGTGCTGATGAGGCTCTGTGCGCGCCGCGAGCCTTAGCC 180
DB 81 GlnLeuGlnLeuLeuAlaPheLeuLeuAspGlnAlaProGlyAlaAlaGlyLeuAla 100
QY 181 CGGGCCCGCAGCGCGCTTAAGACTCTGTGAGCTGAGAGCGCGCGGCGAGTGGCGAG 240
DB 81 GlnLeuGlnLeuLeuAlaPheLeuLeuAspGlnAlaProGlyAlaAlaGlyLeuAla 100
QY 101 ArgAlaArgSerGlyLeuGlnLeuLeuGlnLeuGlnValArgArgGlyGlnCysAspGln 120
DB 241 AGCAACCTGCGGCTGCGGGGCAACTCTCTGCGGTGCGCCGACGACCTGCTGGCG 300
QY 121 SerAsnLeuArgLeuGlnGlnLeuLeuArgValLeuAlaArgHisAspLeuPro 140
QY 301 CACCTGCGCGCAGCGCGCCGCGCAGTGTCTCCAGAACCTATAGCTATGGCACCTCC 360
DB 141 HisLeuAlaArgLysArgArgArgProValSerProGlnArgTyrSerTyrGlnThrSer 160

QY 361 AGCTCTTAAAGAGACAGAGGGTGTGCTGCGCCCTGCGCAGTCAAGCATTTCTGCA 420
Db 161 SerserSerLySArgThrGluGlySerCysArgArgArgGlnSerSerSerSerAla 180
QY 421 AATTCTGACGAGGGTGTGAGGAGCAGGCTCCGCCCAACCAACGGCAGCGGGGAGT 480
Db 181 AsnSerGlnGlnGlyGlnTrpGlnThrGlySerProProThrLySArgGlnArgArgSer 200
QY 481 CGGGCCCGCCAGTGTGTGTGTCAGACGGCGCGCGAGAGGGGCCCGACGCCACCCAG 540
Db 201 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 220
QY 541 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAGCAAGTACCTGTGACATCCGCGTC 600
Db 221 GlnGlnSerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeu 240
QY 601 CGGGTTCAGACAGTACTGTGAGCATGGCCAGGCTTGGACGAGGGGTGGCATCCCG 660
Db 241 ArgValArgAlaGluTrpCysGlnHisGlyProAlaLeuGlnGlnGlyValAlaSerArg 260
QY 661 CGGCCCCAGCGCTGGCGCGAGCTGAGACGTTGTGGAGCAGCGCAGTGTGCGC 720
Db 261 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 280
QY 721 TCAAGGACCTGGGCTGTGTGTGTGTGACATCAAGTTCTGAGACCTTCTATCTGAG 780
Db 281 SerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrLeuAsp 300
QY 781 GCGTTCTGGGGCGAATTACTGAGTGGCGCCCTGCTGCGAGGCCCTCGGGGCGTGTCC 838
Db 301 AlaPheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGln-ProCysGlyAlaCysSer 319

RESULT 11
AAB1551
ID AAB1551 standard; protein; 304 AA.
XX AAB1551;
AC 28-FEB-2001 (first entry)
DT 28-FEB-2001 (first entry)
XX

Db Apoptosis related protein encoded by gene 1 clone H1DOK36.
XX
XX Cytostatic; antiinflammatory; immunosuppressive; antisclerotic; cardiant;
XX
XX vincristin; anti-AIDS; vasotropic; anti-ischemic; antiparkinsonian;
XX
XX anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;
XX
XX colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;
XX
XX autoimmune disorder; multiple sclerosis; viral infection.
XX
XX Homo sapiens.
XX
XX MO200056752-A2.
XX
XX 28-SEP-2000.
XX
XX 15-MAR-2000; 2000WO-US006642.
XX
XX 24-MAR-1999; 99US-0126018P.
XX
XX 17-JUN-1999; 99US-0139638P.
XX
XX 18-AUG-1999; 99US-0149449P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J, Young PA;
XX
XX WPI; 2000-587660/55.
XX
XX N-PSDB; AAA95790.
XX
XX Nucleic acids encoding human apoptosis associated protein, useful for the
XX
XX prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's
XX
XX disease, inflammation and ischemic injury.
XX
XX Claim 11; Page 252-253; 273pp; English.
XX

CC The invention relates to the isolation of genes encoding 9 human
CC apoptosis-related proteins. The nucleotide sequences AAB95790-A95798
CC encode the human apoptosis related proteins AAB1551-B15559. The genes
CC can be used to generate fusion proteins by linking to the gene for the
CC human immunoglobulin G Fc (IgG Fc) portion (AAB95799) for increasing the
CC stability of the fusion protein as compared to the human protein only.
CC The gene and encoded protein may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate apoptosis associated
CC protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,
CC melanomas and lymphomas), inflammation, autoimmune disorders (e.g.
CC multiple sclerosis) and viral infections (e.g. herpes)

XX
XX SQ Sequence 304 AA;
XX

Alignment Scores:
Pred. No.: 2,95e-99 Length: 304
Score: 1413.00 Matches: 280
Percent Similarity: 99.64% Conservative: 0
Best Local Similarity: 99.64% Mismatches: 1
Query Match: 82.34% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-1 (1-909) x AAB1551 (1-304)

QY 67 ATGCTGTGCTTCAACCGTATGTTGAGGTGGGCGGGAACCTGACCGAGTGGAGCTG 126
Db 1 MetLeuSerLeuHisArgMetPheGlnValValGlyGlnLeuThrGluCysGlnLeu 20
QY 127 GAGCTCTGGGCTTTTCTGCTGAATGAGGCTCTTGGCGCCCGGAGGCTTAGCCCGGGCC 186
Db 21 GluLeuLeuAlaPheLeuLeuAspGlnAlaProGlyAlaAlaGlyGlyLeuAlaArgAla 40
QY 187 CGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGGAGCAGTGTGGCGAGAGCAAC 246
Db 41 ArgSerGlyLeuGlnLeuLeuLeuGlnLeuGlnThrGlyGlnCysAspGlnSerAsn 60
QY 247 CTGGCGCTGTGGGCAACTCTCGCGCTGTGCGCCGCAAGACTGCTCCGACCTG 306
Db 61 LeuArgLeuLeuGlnGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuProHisLeu 80
QY 307 GCGGCAAGCGGCGCGCGCAGTGTCTTCAGAGGCTTACTATGCACTTCCAGCTCT 366
Db 81 AlaArgLysArgArgArgProValSerProGlnArgTyrSerTyrGlyThrSerSerSer 100
QY 367 TCAAGGAGCAGAGGGGTAGCTGCGCGCTGCGCGCAGTGAAGAGTTCGAAATCT 426
Db 101 SerLysArgGlnGlnGlySerCysArgArgArgArgGlnSerSerSerSerAlaAsnSer 120
QY 427 CAGCAGGTCAGTGGAGACAGGCTCCGCCCAACCAAGCGCAGCGCGAGTCTGGGCG 486
Db 121 GlnGlnGlnGlnTrpGlnThrGlySerProProThrLysArgGlnArgArgSerArgGly 140
QY 487 CGGCCCCAGTGTGTGTCAGACGGCGCGAGAGGCGGCCCGCAGCCGACAGCAG 546
Db 141 ArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGlnGlnGln 160
QY 547 TCAGACCCCGCAGACCTTCTCTGAAGGCAAGTGAACCTGTGACATCCGGCTT 606
Db 161 SerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArgLeuArgVal 180
QY 607 CGACGAGATCTGCTGAGCATGAGCGCAGCTTGGAGCAGGCGCTGTGCTCCCGCGGCC 666
Db 181 ArgAlaGluTrpCysGlnHisGlyProAlaLeuGlnGlnGlyValAlaSerArgArgPro 200
QY 667 CAGGCGCTGGCGCGCAGCTGAGACGTCGTTGGGAGCGCAGCGACGAGTGTGCTCAAG 726
Db 201 GlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArgSerArg 220
QY 727 GACCTGGGCTGTGTGTGTGACATCAAGTTCTCAGAGCTTCTATGAGGCGCTTC 786
Db 221 AspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrLeuAspAlaPhe 240
QY 787 TGGGCGCATCACTGAGTGGCGCGCTGCTGAGCGCCCTGCGGGCGTGTCTCTGAC 846

Db 241 TPGGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyAlaPheLeuTrnGlu 260
 QY 847 GCCCTGGAGAGGCTGTGGCCGAGGCTGTTCGCTGTGTCAGTGTGATGAGGCT 906
 Db 261 AlAeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerValAspGlnAla 280
 QY 907 GAC 909
 Db 281 Asp 281
 RESULT 12
 AAM25705
 ID AAM25705 standard, protein; 242 AA.
 AC AAM25705;
 XX
 DT 16-OCT-2001 (first entry)
 DE Human protein sequence SEQ ID NO:1220.
 XX
 KM Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiant; central nervous system; virucide;
 KM anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KM antiaggregant; haemostatic; vulnery; antileuc; osteopathic; eczema;
 KM dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
 KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antinaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KM thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 22-DEC-2000; 2000WO-US035017.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 XX
 PA (HYSE-) HYSEO INC.
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457603/49.
 DR N-PSDB; AAH99646.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 treatment and diagnosis of e.g. cancer, ulcer and HIV infection.
 XX
 PS Claim 20; Page 253; 1217pp; English.
 XX
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and cells
 CC they are expressed in, such as: antiinflammatory; antineumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC antileuc; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders
 XX
 SQ Sequence 242 AA;
 Alignment Scores:
 Pred. No.: 8.85e-67 Length: 242
 Score: 987.00 Matches: 196
 Percent Similarity: 97.52% Conservative: 1
 Best Local Similarity: 97.03% Mismatches: 3
 Query Match: 57.52% Indels: 2
 DB: 4 Gaps: 1
 US-10-030-271-1 (1-909) x AAM25705 (1-242)
 QY 310 CGCAGGCGCGCGCCCA-----GTGTCACAGAAAGCTATAGTACCTCCAGC 363
 Db 18 ArgSerHisArgThrProGlnSerValSerProGlnArgTyrSerTyrGlyThrSerSer 37
 QY 364 TCTTCAAGAAGACAGAGGTAAGTCCGCTGCGGAGTCAAGACGTTTGCMAAT 423
 Db 38 SerSerTyrArgThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 57
 QY 424 TCTCAGCAGGCTCAGTGGAGAGACAGGCTCCCGCCCAACAAAGGCGAGCGCGAGTCCG 483
 Db 58 SerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 77
 QY 484 GGCAGGCG 543
 Db 78 GlyArgProSerSerTyrGlyAlaValArgArgArgArgGlyAlaProAlaAlaProGlnGln 97
 QY 544 CAGTCAGAGCCCGCCAGACCTTCTTGAAGGCAAGTGAATCTGTGACATCCGCTCCG 603
 Db 98 GlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
 QY 604 GTTCGAGCAGAGTACCTCCGAGCATGGCGCCGCTTGGAGAGCGCGCGCGCGCGCGCG 663
 Db 118 ValArgAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 137
 QY 664 CCCAGGCGCTGGCG 723
 Db 138 ProGlnAlaLeuAlaArgGlnLeuAlaPheGlyGlnAlaThrAlaValLeuArgSer 157
 QY 724 AGGACCTGGGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTTATCTGGACGCC 783
 Db 158 ArgAspLeuGlnSerValValCysAspIleLeuPheSerGlnLeuSerTyrLeuAspAla 177
 QY 784 TTCTGGGGGCACTACCTGAGTGGCGCGCTGCTGCAAGCCCTGGCGGGCGCTTCTTCACT 843
 Db 178 PheTrpGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyAlaPheLeuThr 197
 QY 844 GAGGCGCTGGAGAGGCTGGGCGGCGGAGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 903
 Db 198 GlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerValAspGln 217
 QY 904 GCTGAC 909
 Db 218 AlaAsp 219
 RESULT 13
 AAM39805
 ID AAM39805 standard, protein; 277 AA.
 XX
 AC AAM39805;
 XX
 DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 2950.
 XX XX Human; nootropic; immunosuppressant; cyrostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.
 XX OS Homo sapiens.
 XX XX WO200153312-A1.
 PN 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US034263.
 PF 26-DEC-2000; 2000WO-US034263.
 XX 23-DEC-1999; 99US-00471275.
 XX 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HSE-) HXSEQ INC.
 PA Tang YT, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou F, Goodrich R, Drmanac RT;
 XX WPI; 2001-442253/47.
 DR N-PSDB; AAI58961.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 4; SEQ ID NO 2950; 10078bp; English.
 PS The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 XX encoded polypeptides (AAM3642-AAM42213) with nootropic,
 CC immunosuppressant and cyrostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX SQ Sequence 277 AA;
 Alignment Scores:
 Pred. NO.: 2.49e-64 Length: 277
 Score: 955.00 Matches: 204
 Percent Similarity: 75.45% Conservative: 5
 Best Local Similarity: 73.65% Mismatches: 27
 Query Match: 55.65% Indels: 41
 DB: 4 Gaps: 6
 US-10-030-271-1 (1-909) x AAM39605 (1-277)
 QY 1 ATGGCGCTATCCGGGTGACCCCGGCGGCTGCGGAGAGATGAGTCTGGACTAC 60
 OS Homo sapiens.
 XX 1 MetalauserserTherProAlaprocystgTgInuAspGluCysLeuAspTyr 20

QY 61 TACGGATGCTGTGCTTCAACCGATGTTGAGGTGGTGGCGGCACTGACCGAGTGC 120
 DB 21 TyrGlyMetLeuSerLeuHisArgMetPheGluValValGlyGlyInLeuThrGluCys 40
 QY 121 GAGCTGAGACTCTGCGCTTTCTGCTGAGTAGAGCTCTGCGCGCCGAGGCTTACCC 180
 DB 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyLeuSer 60
 QY 181 CGGGCCCGGCGGCGCTCAGAGCTCTGCTGAGGTGAGCGCGCGGCGGCGGCGAG 240
 DB 61 ArgAlaArgSerGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 QY 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGCGCTGCTGCGCGCGGCAAGCTTGGCGG 300
 DB 81 SerLeuLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 301 CACCTGGCGGCGAAGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
 DB 101 HisLeuAlaArgValArgArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
 QY 361 AGCTCTCAAG 420
 DB 121 SerSerSerValArgTyrGlnGlySerCysArgArgArgArgArgArgArgArgArg 140
 QY 421 AATTCTGAGCAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGCGGCGGAGT 480
 DB 141 AsnSerGlnGln-----GlySerProProThrIlyValArgHisArgSer 155
 QY 481 CGGGGCGGCGGCGGCTGCTGCTGCGGAGAGCGGCGGAGAGAGGCGGCGGCGGCGG 540
 DB 156 ArgGlyArgProSerGlyValAlaArgArgArgArgArgArgArgArgArgArgArg 175
 QY 541 CAGAGTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 DB 176 GlnGlnSerGlnProAlaArgProSerSerGlnGlyValThr-ThrThrAsn--Le 194
 QY 601 CGGGTTCGAGCAGAGTACTGCGAGCATGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
 DB 194 uGlyPheSer-----HisLeu 200
 QY 661 CGGCGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
 DB 200 rGlyHisThrSerThrGlyAlaHisGlyAlaCysSerProCysPro----- 215
 QY 721 TCAAGGAGACCTGGGCTGTGTTTGACATCAAGTTCTAGAGCTTCTCT 772
 DB 216 -----HisLeuSerSerGlnProGlyProCysProAl 226
 QY 773 ----ATCTGAGCGCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 817
 DB 226 aSerIleIlyArgProAspSerProHis-----LeuProProCysCys 240
 RESULT 14
 ID AAM41591 standard; protein; 217 AA.
 XX AAM41591;
 AC 22-OCT-2001 (first entry)
 DT Human polypeptide SEQ ID NO 6522.
 DE 22-OCT-2001 (first entry)
 XX 22-OCT-2001 (first entry)
 XX Human; nootropic; immunosuppressant; cyrostatic; gene therapy; cancer;
 KM peripheral nervous system; neuropathy; central nervous system; CNS;
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KM leukaemia.
 XX OS Homo sapiens.
 XX 1 MetalauserserTherProAlaprocystgTgInuAspGluCysLeuAspTyr 20
 PN WO200153312-A1.

XX 26-JUL-2001.
 XX
 PD
 XX
 PF 26-DEC-2000; 2000MO-US034263.
 XX
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehnman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 PI
 XX N-PSDB; AAI60747.
 DR
 XX
 XX WPI; 2001-442253/47.
 PT
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 XX Example 2; SEQ ID NO 6522; 10078bp; English.
 PS
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AA138642-AA142213) with nootropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 XX
 SQ Sequence 217 AA:
 Alignment Scores:
 Pred. No.: 7.37e-63 Length: 217
 Score: 935.50 Matches: 185
 Percent Similarity: 96.39% Conservative: 2
 Best Local Similarity: 95.36% Mismatches: 1
 Query Match: 54.52% Indels: 6
 DB: 4 Gaps: 1
 US-10-030-271-1 (1-909) x AA1591 (1-217)
 QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGTGGAGGAGGATGAGTCTGAGTAC 60
 DB |||||
 Db 30 MetAlaLeuSerGlySerThrProAlaProCysTrpGluGlnAspGluCysLeuAspTyr 49
 QY 61 TACGGAGTGTGCTGCTTCAACCGTATGTTGAGGTGAGGAGCGGGCAATGACCGAGTGC 120
 Db |||||
 Db 50 TyGlyMetLeuSerLeuHisArgMetPheGluValAlaGlyGlnLeuThrGluCys 69
 QY 121 GACCTGAGAGTCTCTGGGCTTTTCTGCTGAGTGAAGAGTCTCTGCGCGCGGAGGTTAGCC 180
 Db |||||
 Db 70 GlnLeuGlnLeuLeuAlaPheLeuLeuAspGlnAlaProGlyAlaAlaGlyGlyLeuSer 89
 QY 181 CGGGCCCGCAGCGGCTAGAGTCTCTGCTGAGAGTGAAGCGCGCGGAGTGGCGCGAG 240
 Db |||||
 Db 90 ArgAlaArgSerGlyLeuGlySerLeuLeuLeuGlnLeuGlnArgArgGlyGlnCysAspGln 109

QY 241 AGCAACCTCGCGGTGCTGGGGCAACTCTCCGCTGCTGGCCCGCGCAGCACTGCTGCCG 300
 Db |||||
 Db 110 SerAsnLeuArgLeuLeuGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 129
 QY 301 CACCTGGCGCAGCAGCGCGCGCCAGTGTCTCCAGAACGATATGACTATGACACTCC 360
 Db |||||
 Db 130 HisLeuAlaArgIlyArgArgArgProValSerProGlnArgIlySerIlyGlyThrSer 149
 QY 361 AGCTTTCAAAAGAGACAAGAGGTTAGCTGCGCTGCGGACGTCAAGCACTTCTGCA 420
 Db |||||
 Db 150 SerSerSerIlyArgThrGlnGlySerCysArgArgArgGlnSerSerSerAla 169
 QY 421 AATTTCACAGAGGTCAAGTGGAGACAGACTCCCGCCCAACGAGGCGAGCGCGGAGT 480
 Db |||||
 Db 170 AsnSerGlnGln-----GlySerProProThrIlyArgGlnArgArgSer 184
 QY 481 CGGGCCCGCGCCAGTGTGCTGCGCAGACGCGCGAGAGGAGGCGCGCAGCCCGCAG 540
 Db |||||
 Db 185 ArgGlyArgProSerIlyGlyAlaArgArgArgArgGlyProGlnProHisProSe 204
 QY 541 CAGCACTCAGAGCGCGCCAGACCTTCTCTGAAGGCAAG 580
 Db |||||
 Db 204 rSerSerGlnSerProProAspLeuProLeuGlySalAlaIys 217
 RESULT 15
 AAB94040
 ID AAB94040 standard; protein; 318 AA.
 XX
 AC AAB94040;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 XX Human protein sequence SEQ ID NO:14195.
 DE
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 PD
 PD 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-00116126.
 PF
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 PA
 PA (HELI-) HELIX RES INST.
 PI Oca T, Isogai T, Nishikawa T, Hayaeshi K, Saito K, Yamamoto J;
 PI Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 DR
 XX
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 14195; 2537bp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX Sequence 318 AA;

Alignment Scores:

Pred. No.:	5,84e-39	Length:	318
Score:	622.50	Matches:	136
Percent Similarity:	63.12%	Conservative:	54
Best Local Similarity:	45.18%	Mismatches:	90
Query Match:	36.28%	Indels:	21
DB:	4	Gaps:	6

US-10-030-271-1 (1-909) x AAB94040 (1-318)

QY 34 TGGAGAGAGATGAGTGCCTGAGTACTAGCGGATGCTGCTTACCGTATGTTGAG 93
DB 12 TTPROGLUGLUNHISGLYLGLNGLUNHISGLYLEUTYRSELEUNHISARGMECPHEASP 31
QY 94 GTGGTGGCGGCGCAACTAGCCAGAGTGCAGCTGAGCTCTGCGCTTCTG---CTGGAT 150
DB 32 ILEVAlGLYTHNHSLEUTHNHSLEUTHNHSARGVALARGVALLEUSERPHELEUPHEVALASP 51
QY 151 GAGGCTCTGCGCGCGCGCGCGAGCTTACCGCGCGCGCGCGAGCGGCTTACGCTCTGCTG 210
DB 52 VALLEASRAPHNHSGLUARGGLYLEU-----ILEARGANGLYARGASPHELEULEU 69
QY 211 GAGCTGGAGCGCGCGCGCGCGAGGAGCAACTGCGGCTGCTGCGGCAACTCTCTG 270
DB 70 ALALEUGLUAARGGLNGLYARGCYASRGLUSERNPHNARGGLNVALLEUGLNULEU 89
QY 271 CGCGTGTGCGCGCGCGCAAGCTGCTGCGCGCAAGCTGCGCGCGCGCGCGCACTG 330
DB 90 ARGILEIETHNARGNHSASRLEUPROTYRVALTHNLEULYARGARGHVALAVAL 109
QY 331 TCTCCA-----GAAGCTATAGCTATGACCTCC-----AGCTTTCAAG 372
DB 110 CYAPROASRLEUVALASRPLYSTYRLEUGLUNHISLEUTHNHSLEUTHNHSLEUTHNHS 129
QY 373 AGGACAGAGGAGTGCCTGCGCGCTGCGGAGTCAAGAGTCTGCAATTCACAGAG 432
DB 130 ALALEUSERAPROGLUNPROARGPROBRODINPROSERLYSTHVALPROPHIATYR 149
QY 433 GGTGAG-----TGGAGACAGAGCTCCCGCCCAACCAAGCGGAGCGGAGTCCGAGC 486
DB 150 PROVALVALYSCYAPROTHSERGLYPROGLINMETCYSESLYARGPROALARGGLY 169
QY 487 CGGCGCACTGCTGCTGCGCAAGCGCGCGGAGAGGGGCGCCAGCGCGCAAGCGAGCAG 546
DB 170 ARGALATHNLEUGLUSERGLNARGLYARGARG----- 180
QY 547 TCAGAGCGCGCGCGAGCTTCTGTAAGGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 606
DB 181 ---LYSPROVALTHNPROASRPROLYSGILNTHNTHNTHNTHNTHNTHNTHNTHNTHNTHN 199
QY 607 CGAGCAGAGTACTGAGCAGTACGCGCGAGCTTGGAGCAGGCGGAGTCCGCGCGCGCGCG 666
DB 200 ARGALAGLUTYRCYAGLNIHISGLIETHNHSLEUGLNIHISLEUVALPHESERANLYSGIN 219
QY 667 CAGGCGCTGCGCGCGCGCGAGCTGTTGGGAGCGCGCGAGCTGCTGCTGCTGCTGCTGCTG 726
DB 220 ASRPROLEUGLUAARGGLNPHNGLUARGPHEANGLINLAENHNLLEULYSESLYARG 239

QY 727 GACCTGGGCTCTGTGCTTTGTGACATCAAGTTCTACAGCTCTCTACTGAGCGCTTC 786
DB 240 ASRLEUGLUSERILEIETHNHSASRPLLEYSRHESESGIULNTHNTHNTHNTHNTHNTHNTHN 259
QY 787 TGGGCGACTACTGAGTGGCGCGCTGCTGCGAGCCCTGCGGCGCGTGTCTTCACTGAG 846
DB 260 TTPARGASRPTYRILEANGLYSERLNLEUGLNUALNLEULYSGILYVALPHNTHNTHNTHN 279
QY 847 GCCCTGGAGAGGCTGTGGCGCGGAGGAGCTGTTGCGCTGCTGCTGCTGCTGCTGCTGCTG 906
DB 280 SERLEULYSGILNLAVALIHLNHSGLUALIHLNHSLEULEUVALNENVALASRGLUGLU 299
QY 907 GAC 909
DB 300 ASP 300

Search completed: March 22, 2005, 01:35:02
Job time : 135.485 secs

THIS PAGE LEFT BLANK

RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Testis;Tercar;Tercar;Tercar;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Matita H.,
 RA Sekine M., Obaishi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Salto K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagatani K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M.,
 RA Shirogaki K., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
 RA Sawagata M., Takahashi K., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuma N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hottu T.,
 RA Kusano Y., Kanohori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togashi S., Komai F., Hara R., Takeuchi K., Arima M.,
 RA Imose N., Muehshino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujikawa T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Kaitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Brain, and Lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strassburg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Krausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Diatchenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Uebin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raba S.S., Logunov N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Gibbs R.A.,
 RA Vallat D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scheraga J., Schein J.E., Jones S.J.M., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP INTERACTIONS WITH CASP8 AND CASP10.
 RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
 RA Alcivar A., Hu S., Tang J., Yang X.,
 RT "DED and DED2 associate with caspase-8/10 and signal cell death.";
 RL Oncogene 22:291-297(2003).
 CC -1- FUNCTION: May play a critical role in death receptor-induced
 CC apoptosis and may target Casp8 and Casp10 to the nucleus. May
 CC regulate degradation of intermediate filaments during apoptosis.
 CC May play a role in the general transcription machinery in the
 CC nucleus and might be an important regulator of the activity of
 CC GTF3C3.
 CC -1- SUBUNIT: Interacts with Casp8, Casp10 and GTF3C3. Homodimerizes
 CC and heterodimerizes with DED.

CC -1- SUBCELLULAR LOCATION: Nuclear, accumulated in subnuclear
 CC structures resembling nucleoli.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=2;
 CC Name=1;
 CC IsoId=Q8MXF8-1; Sequence=displayed;
 CC Name=2;
 CC IsoId=Q8MXF8-2; Sequence=VSP_010312;
 CC -1- TISSUE SPECIFICITY: Expressed in most tissues. High levels were
 CC found in liver, kidney, heart, ovary, spleen, testes, skeletal
 CC muscle and peripheral blood leukocytes. Expression was absent or
 CC low in colon and small intestine. Expression is relatively high in
 CC the tumor cell lines chronic myelogenous leukemia K-562 and the
 CC colorectal adenocarcinoma SW480. Expression is moderate in the
 CC cervical carcinoma HeLa, the Burkitt's lymphoma Raji, the lung
 CC carcinoma A549, and the melanoma G361. In contrast, two leukemia
 CC cell lines, HL-60 (promyelocytic leukemia) and MOL-4
 CC (lymphoblastic leukemia), show relatively low levels.
 CC -1- DOMAIN: Interacts with Casp8 and Casp10 are mediated by the DED
 CC domain.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -1- CAUTION: Ref.5 (AAH13372) sequence differs from that shown due to
 CC a frameshift in position 186.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF443591; AAL48220.1; -;
 CC DR EMBL; AF457575; AAM10835.1; -;
 CC DR EMBL; AY125488; AAM95240.1; -;
 CC DR EMBL; AK075328; BAC11551.1; -;
 CC DR EMBL; BC013372; AAH13372.2; ALT_FRAME.
 CC DR EMBL; BC027930; AAH27930.1; -;
 CC DR EMBL; HGNC:24450; DEDD2.
 CC DR H-InvDB; HIK0015171; -;
 CC DR InterPro; IPR011029; DEATH_like.
 CC DR InterPro; IPR001875; DED.
 CC DR Pfam; PF01335; DED_1.
 CC DR SMART; SM00031; DED; 1.
 CC DR PROSITE; PSS0168; DED; 1.
 CC KW Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
 CC Transcription regulation.
 CC FT DOMAIN 25 104
 CC FT DOMAIN 104 109
 CC FT DOMAIN 155 173
 CC FT VARSPLIC 145 149
 CC FT CONFLICT 27 27 H -> N (in Ref. 2).
 CC FT CONFLICT 56 56 A -> G (in Ref. 2).
 CC FT CONFLICT 79 79 D -> G (in Ref. 4).
 CC FT CONFLICT 207 207 C -> R (in Ref. 2).
 CC FT CONFLICT 230 230 Missing (in Ref. 5; AAH13372).
 CC SQ SEQUENCE 326 AA; 36178 MW; 3F7B0B307CC870CD CRC64;
 CC
 CC Alignment Scores:
 CC Pred. No.: 1.89e-80 Length: 326
 CC Score: 1543.00 Matches: 302
 CC Percent Similarity: 99.67% Conservative: 0
 CC Best Local Similarity: 99.67% Mismatches: 1
 CC Query Match: 89.92% Indels: 0
 CC DB: 1 Gaps: 0
 CC
 CC US-10-030-271-1 (1-909) x DED2_HUMAN (1-326)
 CC
 CC QY 1 ATGGCGTATCCGGGCGACCGCCCGCTGCTGGAGGAGATGATGCTGACTAC 60
 CC DB 1 MetalauserselrlyserltnProlaaProCystrpdlunlunspglucySleuAapTryr 20

QY 61 TACGGATGCTGCTGCTTCAACCGTATGTTGAGAGGTGTGGCGGGCAATGACCGAGGC 120
DB 21 TyGclwleuSerleuHlaIarGMeIphelgluValglYglInleuThGlucys 40
QY 121 GAGTGGAGCTCTGCGCTTCTCTGAGATGAGCTCTGCGCGCGGAGGCTTAAACC 180
DB 41 GluIeuGlIuIeuLualAphelLeuAbspIuAlaProglYAlaAlaGlYglYleuA 60
QY 181 CGGGCCGGCGCGGCTTGAAGCTCTGCTGAGATGAGCGCGCGGCACTGGCGGAG 240
DB 61 ArgAlaIarGserGylYleuGlIuIeuLualAargGlYglInCyAaspGlu 80
QY 241 AGCAACCTGGCGCTGCTGAGGCAACTCTGCGCGCTGAGCGCGCAACCTGCTCCG 300
DB 81 SerIeuLualAargLeuGlYglInleuLualAargHlaAaspLeuPro 100
QY 301 CACTGGCGCGCAACGGCGCGCGCTGCTTCCAGAACGCTATGATGCACTCC 360
DB 101 HleuLualAargYlAargArgArgProValSerProgluArgYlYrSerYglYThrSer 120
QY 361 AGCTCTCAAGAGAGAGAGAGAGAGTGGCGCGCTGCGCGCTGAGAGAGTTCGCA 420
DB 121 SerSerSerYlAargInglYglYSerCyArGArgArgArgInSerSerSerAla 140
QY 421 AATTCTGAGAGAGGCTGAGTGGAGAGAGCTCCCGCCCAACCAAGCGGAGCGGAGT 480
DB 141 AmsBerInglInglYglInThrdYlSerProProThrlYlAargInAargSer 160
QY 481 CGGGCGCGCGCGAGTGTGTGTCCAGACGCGCGGAGAGGGCGCGCAACCGCCAG 540
DB 161 ArgGlYlAargProSerGylYglYAlaArgArgArgArgYlAProAlaProGlIn 180
QY 541 CAGAGTCAGAGCGCGCGAGCTTCTCTGAAGCAAGAGAGCTGAGATCCGCTC 600
DB 181 GlInSerInglYglInProAlaArgProSerSerInglYlYlAThrCyAaspIleArgLeu 200
QY 601 CGGCTTGCAGAGAGTACTGCGAGATGAGCGCGAGCTTGGAGAGCGGCTGAGTCCCG 660
DB 201 ArgValAargHlaIarGylYrCyglInHlaIarProAlaLeuGlInglYlAAlaSerArg 220
QY 661 CGGCGCGAGCGCTGCGCGCGAGCTGAGACGCTGTTGGAGCGCGCCAGCGTCTCGC 720
DB 221 ArgProGlInAlaLeuAlaArgGlInleuAaspValAphelglYlnAlaThrAlaValleuAarg 240
QY 721 TCAAGGAGCTGGGCTGTGTGCTTGTGACATCAAGTTCAGAGCTCTCTAATCTGAC 780
DB 241 SerArgAaspLeuGlYSerValValCyAaspIleYlAaspSerGylLeuSerYlLeuAasp 260
QY 781 GCTTCTGAGGCGAGTACTGAGTGGCGCGCTGCTGAGCGCGCTGCGGAGCGTCTTCG 840
DB 261 AlAphelTrpGlYlAaspYlYrleuSerGylYlAleuLeuGlInAlaLeuAargYlYlAphel 280
QY 841 ACTGAGCGCGCTGCGAGAGCTGTGGCGCGGAGGCTGTCTGCTGCTGCTGAGT 900
DB 281 ThrGlInAlaLeuAargGlInAlaValglYArgGlInAlaValAargLeuLeuValSerValAasp 300
QY 901 GAGGCTGAC 909
DB 301 GluAlaAasp 303

RESULT 2
DED2_MOUSE
ID DED2_MOUSE STANDARD; PRT: 330 AA.
AC 08QZV0; 08QZV1;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA-binding death effector domain-containing protein 2 (DED-containing protein FLAME-3).
GN Name=dedd2; Synonyms=flame3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J;
RX MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T.,
RA Alnemri E.S.;
RT "Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the p115C102 subunit of human transcription factor TIFC.";
RT Cell Death Differ. 9:439-447(2002).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
RT Alcivar A., Hu S., Tang J., Yang X., Yang X.;
RL "DED2 and DED2 associate with caspase-8/10 and signal cell death.";
RL Oncogene 22:291-297(2003).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Retina;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepieton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Ueda T.B., Tomihata S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fealey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalek U., Smalley D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [1]
CC -1- FUNCTION: May play a critical role in death receptor-induced apoptosis and may target CASP8 and CASP10 to the nucleus. May regulate degradation of intermediate filaments during apoptosis. May play a role in the general transcription machinery in the nucleus and might be an important regulator of the activity of GTF3C3.
CC -1- SUBUNIT: Interacts with CASP8, CASP10 and GTF3C3. Homodimerizes and heterodimerizes with DEDD (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear, accumulated in subnuclear structures resembling nucleoli (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=08QZV0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=08QZV0-2; Sequence=VSP_010313; VSP_010314;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Expression is high in liver, heart, kidney, and testis but low in brain, spleen, lung, and skeletal muscle.
CC -1- DOMAIN: Interactions with CASP8 and CASP10 are mediated by the DED domain (By similarity).
CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
CC -----

DR EMBL: AR545756; AAM10936.1; -
 DR EMBL: AR543541; AAM33179.1; -
 DR EMBL: BC037043; AAH37043.1; -
 DR MGD; MGI:1914629; 2410050E1R1K.
 DR InterPro: IPR011029; DEATH_like.
 DR InterPro: IPR001875; DED.
 DR Pfam: PF01335; DED; 1.
 DR SMART; SM00031; DED; 1.
 DR PROSITE; PS50168; DED; 1.
 DR Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
 KM Transcription regulation.
 FT DOMAIN 25 104 DED.
 FT DOMAIN 104 109 Nuclear localization signal (Potential).
 FT DOMAIN 156 174 Bipartite nuclear localization signal
 (Potential).
 FT VARSPDIC 1 171 Missing (in isoform 2).
 FT VARSPDIC 172 200 /FTID=VSP_010313.
 FT VARSPDIC 172 200 RRRAGLAAGQHOHOELGRPSSEGVTC -> MGKQAGRG
 FT EYTCPRPTARAKMSWSLS (in isoform 2).
 FT FTID=VSP_010314.
 FT SEQUENCE 330 AA; 36786 MW; 889BC4F9E01304B0 CRC64;

Alignment Scores:

Pred. No.:	2,52e-72	Length:	330
Score:	1401.00	Matches:	281
Percent Similarity:	93.16%	Conservative:	5
Best Local Similarity:	91.53%	Mismatches:	17
Query Match:	81.64%	Indels:	4
DB:	1	Gaps:	2

US-10-030-271-1 (1-909) x DED2_MOUSE (1-330)

QY 1 ATGGCCCTATCCGGGCTGACCCCGCCGCTGCTGGAGAGATAGTCCCTGACCTAC 60
 Db 1 MetAl1euser1ylerThrProAlaProSerTrpGluuAspGluCysLeuAspTyr 20
 QY 61 TACGGAGTGTGCTGCTTCAACCGTATGTTGAGTGTGCGCGGCAACGACCGAGTGC 120
 Db 21 TyrglYmetleuserleuH1sarGmetPhegluValaGlyglYgluIneuThrgluCys 40
 QY 121 GAGCTGAGCTCCTGCTTCTGCTGATGAGTCTCTGCGCCCGCGGAGGCTTAGCC 180
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaProGlyGlyLeuAla 60
 QY 181 CGGCGCCGACGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGAGTGCAGG 240
 Db 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuArgGlyGlyCysAspGlu 80
 QY 241 AGGAACCTGCGGCTGCGGCGCAACTCCTGCGGCTGCGCGCGGACGACCTGCGCG 300
 Db 81 SerAsnLeuArgLeuLeuSerGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
 QY 301 CACCTGCGCGGACGCGCGCGCGGACGCTCTCCAGAACGCTATAGCTATGACACC--- 357
 Db 101 HisLeuAlaArgYsArgArgArgProValSerProGluArgYrSerTrpGlyAlaAsnPro 120
 QY 358 TCAGGCTTTCAAGAGACAGAGGAGTACGCTGCGGCTGCGCGGAGTCAACGAGTTCT 417
 Db 121 SerSerSerSerYsArgTrgGluAspSerCysArgArgArgArgGlnAlaSerSerSer 140
 QY 418 GCAAATTCTCAGAGGCTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGCGCGCGG 477
 Db 141 SerAspSerProGlnSerGlnTrpAspTrpGlySerProThrTrpYsArgGlnAlaArg 160
 QY 478 AGTCGAGGCGCGGCGGAGTGTGTCAGACGCGGCGGAGAGGGGCCGACGCGACCC 537
 Db 161 SerArgGlyArgProSerSerGlyAlaArgGlnArgArgArgAlaGlyLeuAlaAlaSer 180
 QY 538 -----CAGAGCAGTCAAGCGCGGACGCTTCTCTGAGGCAAAAGTGAAGTGT 588
 Db 181 GlnGlnHisGlnGlnHisGlnGlnGluLeuGlyArgProSerSerGlnGlyValThrCys 200
 QY 589 GACATCCGCGCTCCGGGTTTCAGACAGAGTACTCGAGCATGGGCGACGCTTGGAGCAGGCG 648

Db 201 AspIleArgLeuArgValaArgAlaGlyTrpCysGlnHisGlyProAlaLeuGlnGlnGly 220
 QY 649 GTGGCATCCCGCGGCGCGGCGGCGGCTGCGCGGACGCTGAGACGCTTTGGGACGCGCAC 708
 Db 221 ValAlaSerArgArgProGlnAlaLeuAlaArgGlnLeuAspAlaPheGlyGlnAlaThr 240
 QY 709 GCAGTCTGCGCTCAAGGACCTGGGCTCTGCTGCTTTCGACATCAAGTTCTAGAGCTC 768
 Db 241 AlaValLeuAlaSerArgAspLeuGlySerValaValaCysAspIleYsPheSerGlnLeu 260
 QY 769 TCCTATCTGAGCGCTTCTGCGGCGGCGACTACCTGAGTGGCGGCGCTGCGGCGGCGG 828
 Db 261 SerTrpLeuAspAlaPheTrpGlyAspTrpLeuSerGlyAlaLeuLeuGlnAlaLeuArg 280
 QY 829 GCGGCTTCTGCTGACTGAGGCGGCTGCGGAGGCTGTGGGCGGCGGAGGCTTGGCGCTG 888
 Db 281 GlyValPheLeuThrGlnAlaLeuArgGluAlaValaGlyArgGluAlaValaArgLeuLeu 300
 QY 889 GTGAGTGTGATGAGGCTGAC 909
 Db 301 ValSerValaAspGluAlaAsp 307

RESULT 3

QY 0919M3 PRELIMINARY; PRT; 369 AA.
 ID Q919M3;
 AC Q919M3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE dedd1.
 GN Name=dedd1;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20373792; PubMed=10917738;
 RA Inohara N., Nunez G.;
 RT "Genes with homology to mammalian apoptosis regulators identified in
 RT zebrafish".
 RL Cell Death Differ. 7:509-510(2000).
 DR EMBL; AF232226; AAF66963.1; -
 DR ZFIN; ZDB-GENE-000616-2; dedd1.
 DR GO; GO:000515; F:protein binding; IEA.
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.
 DR InterPro; IPR011029; DEATH_like.
 DR InterPro; IPR001875; DED.
 DR Pfam; PF01335; DED; 1.
 DR PROSITE; PS50168; DED; 1.
 SQ SEQUENCE 369 AA; 42244 MW; 73B09E9E17EEC247 CRC64;

Alignment Scores:

Pred. No.:	1,21e-30	Length:	369
Score:	672.50	Matches:	156
Percent Similarity:	56.14%	Conservative:	36
Best Local Similarity:	45.61%	Mismatches:	91
Query Match:	39.19%	Indels:	59
DB:	2	Gaps:	7

US-10-030-271-1 (1-909) x Q919M3 (1-369)

QY 34 TGGAGAGAGATGAGTGTGACTACTACGGAGTGTGCTGACCGTATGTTGAG 93
 Db 13 TrpGlnGlnThrGlnCysLeuSerTrpTrpGlnThrLeuSerLeuHisGlnIlePheGln 32
 QY 94 GTGGTGGGCGGCACTGACCGAG---TGCAGCTGAGAGCTTCTGCGCTTCTGCTGAT 150
 Db 33 IleValGlySerGlnLeuThrGlnThrCysGlyGlyGlu---ValAlaPheLeuLeuAsp 51
 QY 151 GAGGCT-----CCT 159

```

Db      52 GlnThrTyrProGlyLysHisProLeuAspProGluGlyTrpThrGluAspLeuProPro 71
Qy      160 GGGCGCCGCCGGA-----
Db      72 GlyProAspGlySerProGlnAlaAsnThrProCysProAspGluLeuLysSerTrpGln 91
Qy      172 -----GGCTTAGCCCGCGCCGCGAGCGCG 195
Db      92 ArgMetGlnProGlnLysGluGlyCysSerIleAlaSerArgHisArgProLysArgProLysGly 111
Qy      196 CTAGAGCTCTCTGAGCTGAGAGCCCGCGGAGAGGCGGAGAGCAAGCACTCCGCGCTG 255
Db      112 ValGluLeuLeuLeuGluLeuGluArgGlyTyrLeuSerAspAlaAsnLeuArgPro 131
Qy      256 CTGGGGGAGCTCCGCGCGCGCTGCGCGCGGAGCAAGCTGCGCGCGAG 315
Db      132 LeuLeuGlnLeuLeuAlaGlyIleLeuThrArgHisAspValLeuProPheValSerGlnLys 151
Qy      316 CGGCGCGCGCGAGTGTCTCCAGAGCAAGCTATGATGAGCACTTCAGCTCTTCAAGAGG 375
Db      152 LysArgArgThrValSerProGlnArgLysIleAspTyrProGluValAspPheArg 171
Qy      376 ACAGAGGATAGCTGCCCTCGCGCGAGTCAAGCAAGTTCGAAATTCTCAGCAGGCT 435
Db      172 GlnAspArgGluValGlySerAsnThrAsnIleProSerPheGluAsnThrGlnAspHis 191
Qy      436 CAGTGGGAGACAGGCTCC-----CCCCAACCAAGCGGCGGAGGCTCG 483
Db      192 HistTrpArgAlaGlySerGlySerSerMetThrSerAlaSerAsnAlaArgLysArg 211
Qy      484 GGGCGCGCCAGTGTGTGTGCAGAGCGGCGGAGAGGCGCGGCGGCGGCGGCGGCGGAG 543
Db      212 GlyArgGlyHisHisTrpSerArgLysSerArg-----GlyProProGluIleGlnPro 229
Qy      544 CAGTCAGAGCCCGCGCAGACCTTCTTGAAAGCAAGTACCTGTGACATCCGCGCTCCG 603
Db      230 GlnSerThrPro-----AsnLysValThrCysAspIleArgLeuArg 243
Qy      604 GTTCGAGCAGAGTACTCTGCGAGCATGGGCGGCGGCTTGAGAGCGGCGTGGCATCCCGGCG 663
Db      244 ValArgAlaGluTyrSerGlnHisGluSerIleLeuArgLysGlyPheProSerAspLys 263
Qy      664 CCCGAGCGCTGGCGCGGAGCTGAGCATGTGTTGGAGCGGCGGAGCGGAGCTCGCTCA 723
Db      264 ProGlnProLeuGluValArgGlnPheGluLeuPheSerArgHisAspLeuLeuArgThr 283
Qy      724 AGGAGCTGGAGCTGTGTGTGTGTGACATCAAGTTCAGAGCTTCTCATCTGAGCGCC 783
Db      284 ArgAspLeuGlySerIleValCysAspIleLysPheSerLysLeuThrAsnLeuAspThr 303
Qy      784 TTCTGGGGGACACTTACCTGAGATGGGCGCTGTGAGAGCCCTCGGCGGCGGCTTCTCACT 843
Db      304 PheTrpProAspIleTyrMetSerGlyAlaLeuLeuLysAlaLeuLysGlyValPheIleThr 323
Qy      844 GAGGCGCTGCGAGAGCTGTGTGTGTGTGTGAGAGCTGTGCGCTGTGAGTGTGATGAG 903
Db      324 AspSerLeuIleArgAlaAlaGlyGlnLysIleValArgLeuLeuValSerValAspGln 343
Qy      904 GCTGAC 909
Db      344 TyrAsp 345

```

RESULT 4

```

O6DHN2 PRELIMINARY; PRT; 404 AA.
AC O6DHN2;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Zgc:92202.
GN Name=Zgc:92202;
OS Brachydanio rerio (Zebrafish) (Danio rerio).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Stausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stadelton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boak S.A., McEwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RA EMBL: BC075935; AAH75935.1; -
RL EMBL: BC075935; AAH75935.1; -
DR GO: 0005515; P:Protein binding; IEA.
DR GO: 0042981; P:regulation of apoptosis; IEA.
DR InterPro: IPR011029; DEATH_like.
DR InterPro: IPR001875; DED.
DR Pfam: PF01335; DED. 1.
DR SMART: SM00031; DED. 1.
DR PROSITE: PS50168; DED. 1.
SQ SEQUENCE 404 AA; 4488 MW; 3C949DB8B3B07B81A CRC64;

```

Alignment Scores:

```

Pred. No.: 1,38e-28 Length: 404
Score: 636.50 Matches: 142
Percent Similarity: 62.31% Conservative: 58
Best Local Similarity: 44.24% Mismatches: 78
Query Match: 37.09% Indels: 43
DB: 2 Gaps: 9

```

US-10-030-271-1 (1-909) x O6DHN2 (1-404)

```

Qy      10 TCCGGGTGACCCCGCGCCCGCTG-----TGGAGAGAGATGAGTGC 51
Db      77 SerSerSerThrArgArgProAlaSerGlyArgValGluProTrpProGluGluAlaVal 96
Qy      52 CTGAGCTACTACGGGATGCTGTGCTTACCGGATGTTCAGAGTGTGGCGGCGCACTG 111
Db      97 AspAspAlaTyrGlyLeuTyrSerLeuHisArgMetPheAspIleValGlyAlaGlnLeu 116
Qy      112 ACCGAGTGGCGAGCTCTGCGCTTCTG-----CTGATGAGGCTCCT 159
Db      117 ThrHisArgAspValArgValIleSerPheLeuPheValAspValIleAspGlyTyr--- 135
Qy      160 GGGCGCGCCGAGAGCTTAGCCCGGCGGAGCGGCGCTAGAGCTCTGTGAGCTGAG 219
Db      136 ---GluArgGlyGlyIle-----ArgSerGlyArgAspPheLeuLeuAlaLeuGlu 151
Qy      220 CGCGCGGCGAGTGTGGCGGAGAGCAAGCTGCTGTGGGGCAACTCTCGCGGTCTG 279
Db      152 ArgGlnIlyArgCysAspGluThrAsnPheArgHisValLeuGlnLeuLeuArgGlyIle 171

```



```

CC      in the nucleus and might be an important regulator of the activity
CC      of GPR33. Inhibits DNA transcription in vitro (By similarity).
CC      -1 SUBUNIT: Interacts with CASP8, CASP10, KR18, KR178, CASP3 and
CC      FADD. Homodimerizes and heterodimerizes with DED2.
CC      -1 SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
CC      during C9p5-mediated apoptosis where it is localized in the
CC      nucleoli (By similarity). Following apoptosis induction, the mono
CC      and/or dubiquitination form increases and forms filamentous
CC      structures that colocalize with KR18 and KR178 intermediate
CC      filament network in simple epithelial cells.
CC      -1 ALTERNATIVE PRODUCTS:
CC      Name=1;
CC      IsoId=O75618-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=O75618-2; Sequence=VSP_003846;
CC      Note=No experimental confirmation available;
CC      -1 TISSUE SPECIFICITY: Widely expressed with highest levels in
CC      testis.
CC      -1 PFM: Exists predominantly in a mono- or dubiquitinated form.
CC      -1 SIMILARITY: Contains 1 death effector (DED) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; AF083236; AAC33105.1; -
CC      EMBL; AF100341; AAD16414.1; -
CC      EMBL; AF043733; AAC80280.1; -
CC      EMBL; AJ010973; CAA09445.1; -
CC      EMBL; AF064605; AAC17110.3; -
CC      EMBL; BC016724; AAC17214.1; -
CC      EMBL; BC013910; AAH13910.1; -
CC      Genew; HGNC:2755; DEDD.
CC      H-invDB; HIX0001231; -.
CC      MIM; 606841; -.
CC      DR GO; GO:0005737; C:cytoplasm; ISS.
CC      DR GO; GO:0005770; C:nucleolus; ISS.
CC      DR GO; GO:0003677; P:DNA binding; ISS.
CC      DR GO; GO:0006917; P:induction of apoptosis; ISS.
CC      DR GO; GO:0006625; P:induction of apoptosis via death domain rec. .; TAS.
CC      DR GO; GO:0016491; P:negative regulation of transcription; ISS.
CC      DR GO; GO:0007283; P:spermatogenesis; TAS.
CC      DR InterPro; IPR011029; DEATH_like.
CC      DR InterPro; IPR001875; DED.
CC      DR Pfam; PF01335; DED; 1.
CC      DR SMART; SM00031; DED; 1.
CC      DR PROSITE; PS50168; DED; 1.
CC      KM Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
CC      Repressor; Transcription regulation.
CC      FT DOMAIN 25 103 DED.
CC      FT VASPLIC 194 194 DED.
CC      FT FT D -> GEEIQGFQRMSLEGYSKEILGHMAVVAIQ (in
CC      FT isoform 2).
CC      FT CONFLICT 13 13 /FTId=VSP_003846.
CC      FT SEQUENCE 318 AA; 36794 MW; FP9DSFF961F6BB6 CRC64;
CC
CC      Alignment Scores:
CC      Pred. No.: 2.54e-27 Length: 318
CC      Score: 614.50 Matches: 135
CC      Percent Similarity: 62.79% Conservative: 54
CC      Best Local Similarity: 44.85% Mismatches: 91
CC      Query Match: 35.81% Indels: 21
CC      DB: 1 Gaps: 6
CC
CC      US-10-030-271-1 (1-909) x DEDD_HUMAN (1-318)
CC
CC      34 TGGAGAGAGATGACTGCTGGACTACTACGAGATGCTGCTGCTTACACCGATGTCGAG 93
CC      ||| |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::

```

[illegible]

Accession	Gene	Protein	Length	Score	Percent Similarity	Best Local Similarity	Query Match	DB
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.							
CC	NCBI_TaxID=10116;							
CC	SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.							
CC	TISSUE=Testis;							
CC	MEDLINE=99049260; PubMed=9632420; DOI=10.1210/en.139.12.4839;							
CC	Lee C.P., Hsu S.Y., McGee E.A., Salanova M., Hsueh A.J.W.;							
CC	"DEFT, a novel death effector domain-containing molecule predominantly							
CC	expressed in testicular germ cells.";							
CC	Endocrinology 139:4839-4848(1998)."							
CC	-1- FUNCTION: A scaffold protein that directs CASP3 to certain							
CC	substrates and facilitates their ordered degradation during							
CC	apoptosis. May also play a role in mediating CASP3 cleavage of							
CC	KRT18. Regulates degradation of intermediate filaments during							
CC	apoptosis. May play a role in the general transcription machinery							
CC	in the nucleus and might be an important regulator of the activity							
CC	of GTR3C3. Inhibits DNA transcription in vitro (By similarity).							
CC	-1- SUBUNIT: Interacts with CASP8, CASP10, KRT18, KRT18, CASP3 and							
CC	FADD. Homodimerizes and heterodimerizes with DED2 (By							
CC	similarity)".							
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus							
CC	during CD95-mediated apoptosis where it is localized in the							
CC	nucleoli. Following apoptosis induction, the mono and/or							
CC	oligubiquitination form increases and forms filamentous structures							
CC	that colocalize with KRT8 and KRT18 intermediate filament network							
CC	in simple epithelial cells (By similarity)".							
CC	-1- TISSUE SPECIFICITY: Widely expressed with highest levels in							
CC	testis. Within the testis, highly expressed in germ cells but not							
CC	expressed in Sertoli cells.							
CC	-1- DEVELOPMENTAL STAGE: First detected in 20-day-old animals. Reaches							
CC	a peak at 30 days.							
CC	-1- PTM: Exists predominantly in a mono- or diubiquitinated form.							
CC	-1- SIMILARITY: Contains 1 death effector (DED) domain.							
CC	-----							
CC	CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration							
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -							
CC	CC the European Bioinformatics Institute. There are no restrictions on							
CC	CC use by non-profit institutions as long as its content is in no way							
CC	CC modified and this statement is not removed. Usage by and for commercial							
CC	CC entities requires a license agreement (See http://www.isdb-sib.ch/announce/bsd-ch).							
CC	CC or send an email to license@sib-sib.ch).							
CC	-----							
CC	EMBL; AF053362; AAC80287.1; -							
CC	DR GO; GO:0005737; C:cytoplasm; ISS.							
CC	DR GO; GO:0005730; C:nucleolus; ISS.							
CC	DR GO; GO:0003677; F:DNA binding; ISS.							
CC	DR GO; GO:0006917; P:induction of apoptosis; ISS.							
CC	DR GO; GO:0016481; P:negative regulation of transcription; ISS.							
CC	DR InterPro; IPR011029; DEATH_like.							
CC	DR InterPro; IPR001875; DED.							
CC	DR Pfam; PF01335; DED; 1.							
CC	DR SMART; SM00031; DED; 1.							
CC	DR PROSITE; PS0168; DED; 1.							
CC	CC Apoptosis; DNA-binding; Nuclear protein; Repressor;							
CC	KW Transcription regulation.							
CC	FT DOMAIN							
CC	FT 25 103 DED.							
CC	SEQUENCE 318 AA; 36847 MW; B8751791F66A03DE CRC64;							
CC	Alignment Scores:							
CC	Pred. No.:	4.3e-27	Length:	318				
CC	Score:	610.50	Matches:	131				
CC	Percent Similarity:	62.13%	Conservative:	56				
CC	Best Local Similarity:	43.52%	Mismatches:	93				
CC	Query Match:	35.58%	Indels:	21				
CC	DB:	1	Gaps:	5				

QY	94	GTGGGGGGGGGAACTACCCAGAGGCGAGGAGCTCTGGCTTTCTG---CTGAT	150
Db	32	IIeValGIyThrHIsleuThrHIsArgSerValAlaGValIleuSerPheIleuHeValAsp	51
QY	151	GAGGCTCTGGCGCGCCGAGAGCTTAGCCCGGCGCCCGAGCGGCTTAGAGCTCTGGCTG	210
Db	52	ValIIeaSerPheNIIeGIuArgGIyLeu-----IIeaGAsnGIyARgAsPheIleuLeu	69
QY	211	GAGCTGAGAGCGCGGGGAGTGGCGGGGAGAGCAACCTGGCTGTCTGGGAGCACTCTGG	270
Db	70	AlaIeuGIuArgGInGIyArgCyAspSerIleuSerAsnPheArgGInValIleuGInIleuLeu	89
QY	271	CGCGTGTGAGCCCGCGAGCACTCTGGCTGGCGAGCTCTGGCGGCGAGCGCGGCCAGTG	330
Db	90	ArgIIeIIeThrThrArgHIsAspIleuIeuProTyValIIThrIleuTyIAsArgHIsAlaVal	109
QY	331	TCGTCCA-----GAACGCTATAGTACATGAGCACTACAGCTCTTCAAGAGACAGAG	381
Db	110	CysProAspIleuValAspIlySerTyIeuGIuGInuThrSerIIeArgTyValIIThrProArg	129
QY	382	GGTAGCTGCGCTCGCGCTCGGACAGCAAGCACTTTCGCAATTCACAGAGGATCGAG	441
Db	130	AlaIeuSerAspProGIuIuProArgProGInProGInProSerIyThrValIProProHIsTyI	149
QY	442	GAGACAGGCTCCCCCAACCAAG-----CGGACGCGCGAGTGGGGC	486
Db	150	ProValIuCyAspProThrSerIySerGInMetCySerIyAsArgProIIaArgGIy	169
QY	487	CGGCGACATGGTGGTGGCGAGACGCGCGGAGAGAGGGCCCGCCAGCGACCCAGCAGCAG	546
Db	170	ArgThrThrIleuGIySerGInArgIyAsArg-----	180
QY	547	TCAGAGCCCGCGACACTTCTCTGAGAGCAAAAGTACCTGTGACATCCGCTCCGGGTT	606
Db	181	--LysSerValIIThrProAspProIySerGIuGInThrCyAspIIeArgIleuArgVal	199
QY	607	CGAGCAGAGTACTGGAGCATGGGCGAGGCTTGGAGCAGAGGGGTGGACATCCGCGCGGCC	666
Db	200	ArgIaGIuTyIuCyAspGInHIsGIuThrAlaIeuGInGIyAsnValIlePheSerAsnIyGIn	219
QY	667	CAGGCGCTGGCGCGGCGAGCTGACGTGTGTGGCAGGCGCCAGCGAGTGTGCTGCTCAAG	726
Db	220	AspProIeuGIuArgGInPheGIuIuArgPheAsnGInIIaIAsnThrIIeIeuIySerArg	239
QY	727	GACCTGGGCTGTGGTGTGTGACATCAAGTTCTAGAGCTCTCTATGTGAGCGCTTC	786
Db	240	AspIeuGIySerIIeIIeIleCyAspIleIyAspIleSerGIuIeThrTyIeuAspIaIlePhe	259
QY	787	TGGGCGCACTACTAGTGGCGGCGCTGTGTCAGAGCGCTGCGGGGCGGTGTCTCTGACTAG	846
Db	260	TrpArgAspTyIleIeAsnGIySerIleuIeuGInIIaIeIuIyGIIValPheIIeThrAsp	279
QY	847	GCCCTGCGAGAGGCTGTGGGCGGGAGGCTGTTCGCTGTCTGTCAAGTGGATGAGGCT	906
Db	280	SerIleuIyGInIIaIeValGIyHIsGIuIIaIleIyIleuValAsnValAspGIuGIn	299
QY	907	GAC 909	
Db	300	Asp 300	

```
Dy      34 TGGAGGAGAAGATGAACGTCCCTGCAGTAATAAAGCGATTGTTCGGTTTCAACGGTAGTGTCAGAA   93  
       ||| |::: ::||| |::: ||||| |::: ||||| |:::  
  
Db     12 TTrrrrooiiuguiuuagValGiugiinglunih6[ueufuyfveeluenhiavghmetrieaar 31
```

05 Mus musculus (Mouse).
0C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
0C Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus
0X NCBI_TaxID=10090;

QY	598	GCCCGATGTACAGGCTTCACTTGTGCTCAGAGGAAGTCTGGGGGCTCTGACTGCTGCT	539
Db	179	oProProProalaglyProAlaAla-----	187
QY	538	GGGGTGGCGCTGGGGGCCCCCTCTCCGGCCGCTGTGCACACACACCTGGGGCCCGAC	479
Db	188	-----AcgProThrProAlaProThrAlaThrProThrProValAlaProPr	203
QY	478	TCCGCGCGCTGCCGCTTGGTGGGGGGAGCCGTCTCCACCTGACCTGTGGAATTG	419
Db	203	oProAlaAlaProThrAlaArgProGlySerProAlaProAlaAlaThrProAlaProTh	223
QY	418	CAGAACTGCTTACATCGCGAGCGCGACCGGACTACCTCTGT-----	376
Db	223	rProThrPro---AlaProThrAlaThrProAlaProThrAlaThrProAlaProGlySe	242
QY	375	-----CCTCTTGAAGAGCTGGA-----	358
Db	242	rThrProGlyAlaProProAlaGlyArgProGlyAlaProProProGlyValArgProGl	262
QY	357	-----GGTCCCATAGCTATAGCGTTCTTGAG	322
Db	262	ySerProProAlaAlaGlySerProProAlaProGlyAlaThr-----	276
QY	331	ACACTGGCGGGCGCGGCTTGC-----GGCGAGGTGGCGACGA-----GG	291
Db	277	-----ProAlaProThrThrProAlaProGlyGlyThrAlaThrProProSerGl	294
QY	290	TGCTGGCGGGCGAGACGCGAGAGATTCCCGACAGCGCGAGTGTGCTTCGCGCAC	231
Db	294	yArgProGlyProAla-----SerThrProAlaProGlyAlaAlaThrProAl	310
QY	230	TGCCCGCGCGCTGCAGCTCCAGCAGAGACTTA-----GGCGGTGGCGGCC	183
Db	310	aProThrAlaThrProAlaProGlyGlyAlaLeuThrProProProGlyArgProGlyAl	330
QY	182	CGGCTAAGCCCTCCGGCGCGCCAGAGAGCTCATCCAGACGAAGAAGCCAGAGACTCAGC	123
Db	330	agly-----ProThrProGlyProGlyGlyThrProPr	342
QY	122	TCGCACCTGGTCAGTTGCCCGCCACCA-----CCTCGAACAATACGGTGAAGCAGACG	69
Db	342	oAlaglyAlaProAlaAlaGlyThrProAlaAlaProProGlnAlaGlyGlyLeuProAl	362
QY	68	ATCCCGTAGTACGAGCATTCACTCTCTCCACACAGGGGCGGGGTGACCCGGAT	9
Db	362	aArgProAlaAlaProAlaGlyAlaAlaAlaProSerThrValProGlySerAlaAlaAl	382
QY	8	AGCGCCA 2	
Db	382	athrPro 384	
RESULT 12			
ID	06C708	PRELIMINARY;	PRT; 659 AA.
AC	06C708;		
DT	25-OCT-2004 (TREMBLrel. 28, Created)		
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)		
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)		
DE	Similar to tr Q95JC9 Sus scrofa Basic proline-rich protein.		
GN	ORFNames=YALI0D26191g;		
OS	Yarrowia lipolytica CL1899.		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetes; Dipodascaceae; Yarrowia.		
OX	NCBI_TaxID=284591;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CL1899;		
RG	Genolever;		
RA	Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,		
RA	Goffart N., Frangeul L., Angie M., Anthouard V., Babour A., Barbe V.,		
RA	Barney S., Blanchon S., Beckerich J.M., Beyne E., Bleykasten C.,		

RA	Despinae A., Boyer J., Catcolico L., Confantolieri F., de Darvar A.,
RA	Dejoris L., Fabre E., Fairhead C., Ferry-Dunazet H., Groppi A.,
RA	Hantzave F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA	Kerret A., Koszki R., Lemaire M., Lesur I., Ma L., Muller H.,
RA	Nicaud J.M., Nikolski M., Ozts S., Ozer-Kalogeropoulos O.,
RA	Pellens S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA	Swennen D., Tekla F., Mesolowski-Louvel M., Westhof E., Wirth B.,
RA	Zentou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA	Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weisenbach J.,
RA	Wincker P., Souclet J.L.;
RT	"Genome evolution in yeasts."
RL	Nature 430:35-44(2004).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CLTB99;
RA	Genoscope;
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR	EMBL: CR382130; CAG81510.1; ..
DR	InterPro: IPR003124; WH2.
DR	Pfam: PF02205; WH2; 1.
SO	SEQUENCE 659 AA; 62466 MW; 2CE5088BC9EF0908 CRC64;
Alignment Scores:	
Pred. No.:	0.000662
Score:	205.00
Percent Similarity:	35.69%
Best Local Similarity:	28.90%
Query Match:	12.52%
DB:	2
	Gaps: 21
US-10-030-271-1 (1-909) x Q6C7Q8 (1-659)	
QY	889 CCAGCAGGCGAACAACCTCCCGGCCACAGCCTTCGACGGGCTTCAGGACAACGC 830
DB	270 ProthrhPheglYalAProProProProProProglYglYAlAProAlatYrGlYala--- 288
QY	829 CCGGCGAGGCGCTGCAGCAGGCGGCCACGCTAGCTAGTCCGCCCGAAGGCGTCCAGATAGG 770
DB	289 ProAlaProProthrhProglYthr---SerSerProLYsProProProLYsProAlaLYs 307
QY	769 AGA-----GCTCGAGAATGTGATGTCAACAACAGCAGGCGCAGGTCCC--TTGAGC 719
DB	308 ArgProProAlaLeuLYsPro-----LYsProLYsIlleProthrhProglYleuLYs 324
QY	718 GCAGCAGTGCAGGTGCGCTGCCCAACAGCTGCAGCTGCCGCCGCGCAGGCTGGGGCGCC 659
DB	325 ProAlaValProthrhProglYglhArg---SerValSerProSerProglYAlaPro 343
QY	658 GGGATGCCAGCCCT-----GCTCCAGGCTGGGCCATGCT----- 623
DB	344 ProProProIlleProglYSerValAlaProSerValArgHlAlaProSerGlnSerVal 363
QY	622 -----CCAGAGCTGCTGTCGAACCCGAG----- 598
DB	364 SerSerIlleAlaSerSerVal-SerThProSerThProProProAlaProProAlaPr 383
QY	597 -CCGAGTGTCAAGCTCACTTTGCTTCAGAGGAAGTGTCGGCGGCTCTGACTGTGCT 539
DB	383 oProProAlaProLY----- 388
QY	538 GGGGTGGCGGTGGGCGCCCTCTCGCGCGCGCTGGCACCACCACTGGGCGGCGCCGAC 479
DB	389 -----GlyAlaAlaProProIlleProglYSerAlaAlaProProAlaProPr 404
QY	478 TCCGCGCGTGGCGGTGGTGGGG-----GGAGCGTGTCTCC 440
DB	404 oProAlaProProAlaGlYrPheglYAlaProAlaProProSerPheglYAlaProThrhP 424
QY	439 ACTGACCTGTGTGAATTTGACAGAACTGTGACTGCCGAGCGGCGCAGCTACCT 380
DB	424 o-----ProProAlaProProAlaProSe 432
QY	379 CTGTCTCTTTGAAAGCTGGAGGTGCACATAGCTAAGCGTTCTGGAGACACTGCCCGC 320

```
Db 432 r1a1Pro-----Prol 436
QY 319 GCCGCTTGGCCGCGAGTGCAGGAGGTCGTT-----GGCGGCGCCAGCA 275
Db 436 aPro--ProAlaProAlaProProSerSerGlnProProSerThrProAlaGlyProAla 455
QY 274 CGCGCAGGAGTT-----GCCCGACGACCGCGCGTTCCTCGCGCGCACTGCC 227
Db 456 MetPheGlyAlaProMetProLysSerProAlaAlaAla-----Ser 469
QY 226 CGCGGCGCTCCAGCTCCAGACAGAGCTTAAGCGCGCGCGCTAAGCCTCCGG 167
Db 470 ProGlyAlaProProProProProPro-----GlyAlaAlaAlaProGlyLeuAla----- 486
QY 166 CGCGCGCAGAGGCTCATCCAGCAGAAAGCCGAGGCTCCAGCTCGACTCGTCAATT 107
Db 487 -----ProProAlaProProAlaGlnProProSerProGlyArgProSerGlyAla 503
QY 106 GCCCGCCGACCACT-----CGAACATACGCTGAAGCGACA 71
Db 504 ProProProProProProProProProAlaProProThraSpGlnPheHisSerMetIleLeu 523
QY 70 GCATCCCGTAGTAGTCCAGGCACTCAT-----CCTCTCCGACGACGCGG 26
Db 524 AspAspGlySerSerSerSerGlySerHisGlyAlaProProProProProProSerAlaPro 543
QY 25 CCG-----GGTCGACCCGCGATAGCGCCA 2
Db 544 ProSerAaGlyGlyHisSerHisGlyAlaPro 554
```

RESULT 13

```
Q7UB18 PRELIMINARY; PRT; 496 AA.
ID Q7UB18 AC Q7UB18;
AC 01-OCT-2003 (TREMblrel. 25, Created)
DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Possible N-terminal part of IF-2.
GN Ordered locus names: SYNW0598;
OS Synechococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22925697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brakhamsa B., Larimer F.W., Land M.L., Hauser L.,
  Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
  Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569690; CAE07113.1;
DR GO; GO:0003743; F:translational initiation factor activity; IEA.
DR GO; GO:0006413; P:translational initiation; IEA.
DR InterPro; IPR006847; IF2_N.
DR InterPro; IPR009061; Putativ_DNA_bind.
DR Pfam; PF04760; IF2_N; 1.
KW Complete proteome.
SQ SEQUENCE 496 AA; 49294 MW; FC20BB864F65F286 CRC64;
```

Alignment Scores:

```
Pred. No.: 0.000766 Length: 496
Score: 204.00 Matches: 91
Percent Similarity: 34.16% Conservative: 19
Best Local Similarity: 28.26% Mismatches: 110
Query Match: 12.45% Indels: 102
DB: Gaps: 16
```

US-10-030-271-1 (1-909) x Q7UB18 (1-496)

```
QY 907 CAGCCCTACCCACATGACGAGCGGCAAGAGCTCCGCGCCACAGGCTCTCGAGGG 848
Db 135 LysProValProArg--ProAlaAlaAlaProProArgProGlnPro--AlaLys 152
```

```
QY 847 CCTCAG-----TCAGAAACAGCCCGGAGGCTTCAGCAGAGGCGCCACTAGTAAT 794
Db 153 ProGlnValValSerLysProLysProAlaThrProAlaThrAla--SerAlaProSer 171
QY 793 CGCCCGCAGAGGGTCCAGATAGG---AGAGCTGTGAACCTTGATGTCCAAACGA--- 740
Db 172 LysProThrAlaProProAlaArgProThrValValLysProThrValAlaLysProThr 191
QY 739 -----CAGAGCCAGAGTCCCTGAGCGGACGA 713
Db 192 ValAlaLysProThrValAlaLysProThrValAlaLysProAlaLysProAla 211
QY 712 CTGCGGTGGCTGCCAAACAGCTCCAGCT---GCCCGCCAGCGCTGGG-----GCC 662
Db 212 AlaAlaLysProAlaProAlaArgProAlaProAlaArgProAlaArgProSer 231
QY 661 GCGCGGATGCGACGCGCTGCTCCAGAGCGGCGCCATGCTGCGAGTCTGCTGCAACC 602
Db 232 AlaAspGlnProLysProArgProAlaAlaAlaProSerArgProThr-----Pro 248
QY 601 GAGCCGATGTCACAGTCACTTTCCTTCAGAGGAGAGGTGCGGCTCTGACTGCT 542
Db 249 GlyAlaGlyGlnLysProGlnIleVal----- 257
QY 541 GCTGGGGTGGCGTGGGCGCCCTCTCCGCGCCCTGTGGCACCACCACTGGGCGGCC 482
Db 258 -----SerArgPro 260
QY 481 GACTCGCGCGCTGCGCTTGTGGGGGGAGCGTGTCCCACTGACCTGTGAGAAAT 422
Db 261 GlySerAlaProArgProGly-----AlaProThrArgProGly----- 273
QY 421 TTGCAGACTGTGACTGCTCCGACGCGCAGCAGCACTACCTGTCTTTGAAGAGC 362
Db 273 ----- 273
QY 361 TGGAGTGCCATAGCTATAGCTTCTGAGACACTGCGCGCGCTTGGCGCCAGGT 302
Db 274 -----AlaProAlaProAlaArgProGly 281
QY 301 GCGGCGCAGAGTGTGCGGCGGCGCAGCAGCGCAGAGTGGCCACAGCGCGAGGT--- 245
Db 282 AlaProValLysAlaGlyProProThrArgProThrProArgProGlnLeuValGlyLys 301
QY 244 ---TGCTCTCGCGCAGCTCCGCGCGCTCCAGCTCCAGCAGAGAGCTTAGGCCCTGC 188
Db 302 ProValProArgArgProGlyThrGlyAlaProThrArgSerGlyAlaGlyAlaProGln 321
QY 187 GGGCCCGGCTAAGCTCCGCGCGCGCGCAG--GAGCTTCATCCAGCAGAAAGCCAGCA 131
Db 322 ArgProGlyThrGlyAlaProGlnArgProGlySerProGlyArgProThrArgProGly 341
QY 130 GCTCCAGCTCCG-----ACTGGTCACTGCCCGCCAGCACTCCGACACTAC 83
Db 342 AlaProAlaArgSerGlyGlyLeuThrLeuGlnValGlyLysProLysArgAsp 361
QY 82 GGTGAAGCAGCAGCATCCGCTAGTAGTCCAGGCACTATCTCTCCAGCAGGCGCG 23
Db 362 GlySerSerThrGlySer-----GlyGlyArgProAlaProProThrArgPro 377
QY 22 GGGTCG 17
Db 378 GlyAla 379
```

RESULT 14

```
IE18_PRIVIF STANDARD; PRT; 1461 AA.
ID IE18_PRIVIF AC P11675;
AC 01-OCT-1989 (Rel. 12, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Immediate-early protein IE180.
```

```

NS Name=IE
OS Pseudocrabies virus (strain Indiana-Funkhouser / Becker) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirinae.
OX NCBI_TaxID=31523;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69315207; PubMed=2546124;
RA Cheung A.K.;
RT "DNA nucleotide sequence analysis of the immediate-early gene of
RL Pseudocrabies virus.";
RL Nucleic Acids Res. 17:4637-4646 (1989) .
RN [2]
RP REVISIONS.
RA Cheung A.K.;
RL Submitted (NOV-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This IE protein is a multifunctional protein capable of
CC migrating to the nucleus, binding to DNA, trans-activating other
CC viral genes, and autoregulating its own synthesis.
CC -1- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -1- PFM: A long stretch of serine residues may be a major site of
CC phosphorylation.
CC -1- SIMILARITY: Belongs to the herpesviruses ICP4/IE140/IE180 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce,
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X15120; CAJ33214.1; -.
CC DR PIR; S04713; EDBBIF.
CC DR InterPro; IPR005205; Herpes_ICP4_C.
CC DR InterPro; IPR005206; Herpes_ICP4_N.
CC DR Pfam; PF03585; Herpes_ICP4_C; 1.
CC DR Pfam; PF03584; Herpes_ICP4_N; 1.
CC KM DNA-binding; Early protein; Nuclear protein; Phosphorylation;
CC KW Trans-acting factor; Transcription regulation.
CC FT DOMAIN 390 405 Poly-Ser.
CC FT FT 958 966 Poly-Ser.
CC SQ SEQUENCE 1461 AA; 149833 MW; 7F31E7ABE403B208 CRC64;

```

[illegible]


```

Db      260 -----ProGlyMetLysGlyHisArgGlyPheAspGlyArgAsnGlyGlyLysGlyGlu 277
QY      123 GCTGAGCTCTGCTGCTCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTGACCGC 182
Db      278 ThrGlyAlaProGlyLeuLys---GlyGluAsnGlyLeuProGlyGluAsnGlyAlaPro 296
QY      183 GCGCCGACGCGCGCTTGA---GCTCCTGCTGAGAGCTGAGCGCGCGCGGACGTGCGGCA 239
Db      297 GlyProMetGlyProArgGlyAlaProGlyGlyLysArgGlyArgProGlyLeu----- 313
QY      240 GACCAACCTGCGCTGCTGCGGCA-----ACT 266
Db      314 -----ProGlyAlaAlaGlyAlaArgGlyAsnAspGlyAlaArgGlySerAspGlyGln 331
QY      267 CCTGCGCTGCTGCTGCGCGCGCGACGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 320
Db      332 ProGlyProProGlyProProGlyThrAlaGlyPheProGlySerProGlyAlaLysGly 351
QY      321 -----CCGCGCAGTGTCTCCAGAACGCTATGCTATGACGCTCCAGCTCTTCAAA 371
Db      352 GluValGlyProAlaGly-SerProGly-----SerAsnGlyAla-ProGly---GlnA 368
QY      372 GAGGACAGAGGAGGAGCTGCGCGCGCGCGCGAGTTCGCAATTCTCAGCA 431
Db      368 TgGlyGlu-----ProGlyProGlyGly----- 375
QY      432 GAGTCACTGAGAGAGAGCTCCCCCGCAACCAAGCGC-----AGCGCGGAG 479
Db      376 --HisAlaGlyAlaGlnGlyProProGlyProProGlyIleAsnGlySerProGlyGlyL 395
QY      480 TCGGG---GCCGCGCCAGTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 524
Db      395 TgGlyGluMetGlyProAlaGlyIleProGlyAlaProGlyLeuMetGlyAlaArgGlyP 415
QY      525 CCCAGCGCGACCCCGACGAGAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 584
Db      415 ro---ProGlyProAlaGlyAlaAsnGlyAlaProGlyLeu----- 427
QY      585 CTGTGACATCCGCTCCGCGGTTCGAGCAGAGTACTGCGAGCATGCGCGCGCGCGCGCGCG 644
Db      428 -----ArgGlyGlyAlaGlyGluProGlyLysAsnGlyAlaLysGlyGluProGlyPro 446
QY      645 GG-----GCTGCGCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 689
Db      446 TgGlyGluArgGlyGluAlaLysGlyIleProGlyAlaProGlyAlaLysGlyGluAspGly 465
QY      690 CGTGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 749
Db      466 LysAspGlySerProGlyGluProGlyAlaAsnGlyLeuProGlyAlaLysGly----- 483
QY      750 CATCAAGTTCAGAGCTCTCTATCTGAGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCG 809
Db      484 -----GluArgGlyAlaProGlyPheArg 491
QY      810 CTTGCTGACAGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863
Db      492 GlyProAlaGlyProAlaGlyIleProGlyGlyLysGlyProAlaGlyGluArgGly--- 510
QY      864 GGGCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 899
Db      511 AlaProGlyProAlaGlyProArgGlyAlaLysGly 522

```

RESULT 3

Q08B3
BHLFI protein - human herpesvirus 4 (strain B95-8)
C/Species: human herpesvirus 4, Epstein-Barr virus
C/Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C/Accession: A03742
R/Banker, A.T.; Deininger, P.L.; Farrell, P.J.; Barrall, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A/Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A/Reference number: A93065; MUID:85035713; PMID:6092825

```

A/Accession: A03742
A/Molecule type: DNA
A/Residues: 1-660 <BAN>
A/Cross-references: UNIPROT:P03181
R/Barr, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; He
Nature 310, 207-211, 1984
A/Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A/Reference number: A03794; MUID:84270667; PMID:6087149
A/Contents: annotation; protein coding region
C/Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-522;
C/Superfamily: human herpesvirus 4 BHLFI protein

Alignment Scores:
Pred. No.: 7.23e-05 Length: 660
Score: 196.50 Matches: 103
Percent Similarity: 32.98% Conservative: 23
Best Local Similarity: 26.96% Mismatches: 136
Query Match: 12.00% Indels: 120
DB: 1 Gaps: 16

US-10-030-271-1 (1-909) x Q08B3 (1-660)
QY      901 CATCCACACTGACACGAGCGGAACAGCTCCGCGCA-----CAGCTCTCGCA 851
Db      169 HisProProProGlyAlaGlyGlnArgProSerGlyProThrGlyGlyArgProAlaAla 188
QY      850 GGGCTCACTGACGAAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 809
Db      189 ProGlyAlaProGlyThr---ProAlaAlaProGlyProGlyGlyAlaAlaValPro 207
QY      808 -----CGCCACTCAGTATGTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 767
Db      208 SerGlyAlaThrProHisProGlyArgGlySerGlyProAlaAspProAlaAlaAla 227
QY      766 GCTGTGAACACTGATGATCAACAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 707
Db      228 ArgLeuProProGlyArgGlnGluProGlyLeuProGlyAlaAspLeuAlaAlaAlaGlnArg 247
QY      706 TGGCTGCGCGCAACAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 647
Db      248 CysProAlaGlyProProProThrArgSerGlyAlaAlaAlaGlnArgThrHisArgArg 267
QY      646 CCT-----GCTCAAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 593
Db      268 ProProGlyCysProArgSerAlaArg-----AsnProGly 279
QY      592 TGTCAAGGCTGATGCTTGCCTTCAGAGAAAGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 539
Db      280 CysProArgThr-----TThrArgArgSerGlyAlaGlnArg 292
QY      538 -----GGGTGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 494
Db      293 GlyHisProProProGly-AlaGlyGlnArgProSerGlyProThrGlyGlyArgProAl 312
QY      493 TGGGCGG---GCCCGG---ACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 440
Db      312 AlaProGlyAlaProGlyThrProAlaAlaProGlyProGlyGlyAlaAlaValPr 332
QY      439 A----- 439
Db      332 osGlyAlaThrProHisProGlyArgGlySerGlyProAlaAspProAlaAlaAla 352
QY      438 -----CTGACCTGCTGAGAA 422
Db      352 ArgLeuProProGlyArgGlnGluProArgLeuProGlnAlaAspLeuAlaAlaGlnArg 372
QY      421 TTGCAGAACTGCTTGACTGCCGACGCGGACGAGCTAACCTC- 379
Db      372 GCGsProAlaGlyProProProThrArgSerGlyAlaAlaAlaGlnArgThrHisArgAr 392
QY      378 -----TGTCTCTTGAAGAGCTGGAGG 356
Db      392 gProProGlyCysProArgSerAlaArgAsnProGlyCysPro----- 406

```

QY 355 TGCATAGCTATAGCTTTGGAGACACTGGCCGGCGCTTGCGGCCAGTGGCA 296
 Db 407 -----ArglnTrp-ArgArgArgSerGlyAlaAlaInArgln 419
 QY 295 GCAGGTGTGGCGGCCAGACAGCGGAGAGTTGCCAGACGCCAGTTGCTTCCG 236
 Db 419 IspProProGlyAlaGlyGlnArgProSerGlyProGlyAlaArgProAla 439
 QY 235 CGCACTGCGCGCGCGCTCCAGACAGAGCTTAGCGCGCTGCGCGCGCGCTA 176
 Db 439 roGlyAlaPro-GlylnProAlaAlaProGlyPro-----GlyGlyAla 454
 QY 175 AGCCTCGCGCGCGCGCGCTCATCCAGACAGAAAGCCAGAGCTCCAGCTCGCACT 116
 Db 455 AlValAlProSerGlyAlaThrProHisPro---GlnArgGlySerGlyProAlaAspPro 473
 QY 115 CGGTCACTGGCCCGC---CCAGCA-----CCTGCAACATACGGTGAACGACA 71
 Db 474 ProAlaAlaAlaArgLeuProProGlnArglnGlnArgLeuProGlnAspLeuAla 493
 QY 70 GCATCCCGTAGTACTCAGGACATCATCTCCAGACAGGGCGCGCGCTGACCCGG 11
 Db 494 AlaAlaGlnArgCysProAlaGlyProProProThrArgSerGlyAlaAlaGlnArg 513

RESULT 4

A45344
 immediate-early protein - suid herpesvirus 1 (strain Kaplan)
 C:Species: suid herpesvirus 1
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
 C/Accession: A45344
 R/Vicsek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.
 Virology 179, 365-377, 1990
 A>Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented ORF
 A/Reference number: A45344; MUID:91021039; PMID:2171211
 A/Accession: A45344
 A/Status: translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1446 <VLC>
 A/Cross-references: UNIPROT:P33479; GB:M34651; NID:G334070; PID:AA47470.1; PID:G334071
 C/Superfamily: herpesvirus immediate-early protein IE175
 C/Keywords: DNA binding; early protein; transcription regulation

Alignment Scores:

Pred. No.: 6.83e-05 Length: 1446
 Score: 196.00 Matches: 108
 Percent Similarity: 33.77% Conservative: 21
 Best Local Similarity: 28.27% Mismatches: 115
 Query Match: 11.97% Indels: 138
 Gaps: 18

US-10-030-271-1 (1-909) x A45344 (1-1446)

QY 886 GCAGGCGAAGAGCTCCCGGCCACAGCCTTCGACAGGCGCTCAGTACAGAAACGCCCC 827
 Db 143 SerGlyProArgProProthProAlaAlaLeuAlaAlaGlnAlaGlyAla---Pro 161
 QY 826 GCAGGCGCTGACAGAGGCGCGCACTCAGTAGTGGCCCGCAAGAGCGTCCAGATAGAGA 767
 Db 162 GlyGlyProGlyArgSerSerProSerAlaAlaSerPro----- 174
 QY 766 GCTCTGAAACTTGATGTCACAAACACAGACAGCCAGAGCTTCCTTAGCGCAGCACTCGG 707
 Db 175 -----AlaSerSerSerGlySerProGlyProSerAlaAlaProArgArg 189
 QY 706 TGG---CCTGCCCAACAGCTCCAGCTGCCCGCCAGCGCTGGGGCCGCCGGATGCCA 650
 Db 190 TlperProAlaArgGlyAspProValGlyGlnProGlyProAlaAla---ArgProArgT 209
 QY 649 CGCCCTGCTCAAGCGTGGCCGACGTCTGCGAGTACTCTGCTGCAACCCGAGCCGAGATGT 590
 Db 209 hrProAlaProProAlaGlnProAlaAlaValAlaAlaAlaProAlaArgArgGlyProA 229

QY 589 CACAGTCACTTTCCTTCAGAGAGAGTGTGGCGGCGCTTGACTGCTGCTGGGCTCGG 530
 Db 229 IAsrProAlaSerProAla-----AlaGlyProValSerAlaProGlyGly 245
 QY 529 CTGGGGCCCTTC---CGCCGCGCTTGGACACCACTGGCGCGCGCC 482
 Db 245 IyGlyAlaProSerAlaGlyGlyAspArgGlyArg-HisHisGlnHisValArgGlyPro 264
 QY 481 -----GACTCCGCGCTGCCGCTTGGTGGG-----GGAGCGCTGT 443
 Db 265 LeuLeuAspGlnProAlaAlaAlaAlaArgArgLeuAspProArgProLeuGlyAlaArgSer 284
 QY 442 CCC---ACTGACCTGTGTAATTCAGAACTGCTTGACTGCTC--- 401
 Db 285 ProValSerSerAsnProAsnSerAsnSerAsnSerThrThrValAlaValGluThr 304
 QY 400 -----GACGCGACGCGACGTAACCTCTGTCTCTTTGAAGAC 362
 Db 305 ValAlaArgGlyProGlnLysAspGlnAspGlyLeu----- 316
 QY 361 TGGAGTGCATAGCTATAGCGTTCTGAGACACTGSCCGGCGCGCTTGCGCGCA--- 305
 Db 317 -----GlyLeuAlaGlyAspGlyGlyAlaProPro 326
 QY 304 -----GTTGCGGACACA-----GGTGTGTC--- 284
 Db 327 GlnArgGlnProArgArgArgAlaGlyGlyGlyAlaLeuArgArgGlyAlaPhe 346
 QY 283 -----GGCCAGACGCGCAGAGATTGCC 260
 Db 347 Ser 366
 QY 259 CCA---GCAGCGGAGGTTGCTTCGCGGACGCTCCGCGCGCTCCAGCTCCAGCAGA 203
 Db 367 ProArgAlaProAlaAlaAlaAlaAlaArgArgSerAlaSerSerSerSerSerSer 386
 QY 202 GCTCT----- 198
 Db 386 erSer 406
 QY 197 -----AGCCGCTTCGGGCGCGGCTTAAGCTTCGGCGGCGCCAGACCTCATCCA 146
 Db 406 rgProGlyAlaProLeuAlaAlaArgAlaGlyProProProSerProProAla-ProAlaAla 425

QY 145 GCAGAAAGGCGAGAGCT-----CCAGCTCGCACT 116
 Db 426 AlaProArgProSerAlaSerSerSerSerSerAlaAlaAlaSerProAlaProAla 445
 QY 115 CGGTCACTTCCCGCCAGCAGCTCGAACAATACGTTGAGAGCAGCAGCATCCGTAGTAGT 56
 Db 446 ProGlnProAlaArgProProArgArg-----LysArgArgSer 458
 QY 55 CCAAGCACTCAT-----CCTCTCCAGACGCGGCGG 23
 Db 459 ThrAsnAsnHisLeuSerLeuMetAlaAspGlyProProProThrArgGlyPro 476

RESULT 5

A61183
 hypothetical protein (sdB region) - Pseudomonas sp.
 C:Species: Pseudomonas sp.
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Nov-1999
 C/Accession: A61183; S27643
 R/Davison, J.; Brunel, F.; Phanopoulos, A.; Prozzi, D.; Terpetra, P.
 Gene 114, 19-24, 1992
 A>Title: Cloning and sequencing of Pseudomonas genes determining sodium dodecyl sulfate l
 A/Reference number: J01118; MUID:92267380; PMID:1587481
 A/Accession: A61183
 A/Molecule type: DNA
 A/Residues: 1-312 <DAV>
 A/Cross-references: EMBL:M86744
 A/Experimental source: ATCC 19151
 C/Comment: A translation in another reading frame of 306 codons from within the nucleotic
 x proteins.

C,Superfamily: proline-rich protein

Alignment Scores:

Pred. No.:	0.000362	Length:	312
Score:	186.00	Matches:	99
Percent Similarity:	32.93%	Conservative:	10
Best Local Similarity:	29.91%	Mismatches:	126
Query Match:	11.36%	Indels:	96
DB:	2	Gaps:	19

US-10-030-271-1 (1-909) x A61183 (1-312)

```

QY 889 CGAGGAGGGAAGAGCTCCCGGCGGACAGACCTCTGACAGGCGCTGACAGGAACAGC 830
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 34 ProIaGlyProIuProGlnHsProGlyPro-----GlyGluuProGlyLeuPro 51
QY 829 CCGGAGGCGCTGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 785
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 52 ProIaGlyProIuProGlyProIaProGlyProIaProGlyProIaProGlyProIa 71
QY 784 AGGCGTCA-----GATAGAGAGCTCTGAGAACTTGATGTCACAAACAGAG--- 737
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 72 ThrArgProIaProIaProIaProIaProIaProIaProIaProIaProIaProIa 91
QY 736 AGGCGGCGCTGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 680
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 92 GlnProGlyGlnArgIaProGlyProGlyProGlyProGlyProGlyProGlyProGly 111
QY 679 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 626
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112 AlaProGlyAlaIaProGlyAlaIaProGlyAlaIaProGlyAlaIaProGlyAla 131
QY 625 GCTGCGAGTACTGCTGCTGCAACCGGAG-----GCCGAGTGCAC 587
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 132 ArgHis-----LeuValGlnProIaProIaProIaProIaProIaProIaProIa 149
QY 586 AGGTCATTGCTGACAGGAGAGTGTGCGGCGGCTGTGCTGCTGCTGCTGCTGCTG 527
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 150 ArgGlyArgGly-----AlaGlyThrAlaArg 158
QY 526 GCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 467
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 159 ArgPro-----AlaAlaGlyGlyHis-----AlaAlaAla 168
QY 466 GCTTGGTGGGCGGAGCTGTCTCCACT-----GACCTGTCTGAGAAAT 422
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 169 GlnLeuAlaArgGlyAspPheLeuProThrGlyAlaProIaAlaGlyProIa----- 186
QY 421 TTGCGAGAGCTGTAAGTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 187 -----ProIaProIaAspGlyGlyArgPro--GlySerLeuProIa 200
QY 361 TGGAGGTGCATAGTATAGCTTCTGAGACACTGGCGGCGGCGGCGGCGGCGGCGGAGT 302
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 201 GlyArg-----HisProGly 205
QY 301 GCGGCGAGAGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 242
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 206 AlaArgArgGlyGlyAlaGlyProGlyAlaValGlyGlyArgProIaAlaGlyHis 225
QY 241 TCTGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 226 ArgMetArgGlnLeuHisGlyAlaGlyGlyAlaGlyGlyArgGlyAlaGlyHis 245
QY 181 GGGCTAGAGCTCCGG-----CGGCGGCGAGGCTCTCCAGAGAA--- 140
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 246 GlyAlaThrGlyArgGlyArgArgArgArgArgArgArgArgArgArgArgArgArg 265
QY 139 -----AGGCGAGAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCC 113
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 266 GlyAlaArgProIaAlaIaThrLeuGlyLeuArgProGlyGlnProIaArgGly----- 283
QY 112 TCGATTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 53

```

Db 284 -----AlaHisProIaThrGlyCysArgGlyAspAla-----ArgAlaAspPro 298

QY 52 GGCATCATCTCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20

Db 299 ArgArgGlyArgProHisProThrGlyLeuGly 309

RESULT 6

CGCH1S collagen alpha 1(I) chain - chicken (tentative sequence) (fragments)

C/Species: Gallus gallus (chicken)

C/Date: 12-Aug-1981 #sequence revision 06-Jul-1982 #text_change 31-Mar-2000

C/Accession: A90458; A90181; A02857

R/Higberger, J.H.; Corbett, C.; Dixit, S.N.; Yu, W.; Seyer, J.M.; Kang, A.H.; Grose, J.

Biochemistry 21, 2048-2055, 1982

A/Title: Amino acid sequence of chick skin collagen alpha1(I)-CB8 and the complete primary

A/Reference number: A90458; PMID:82231995; PMID:7093229

A/Accession: A90458

A/Molecule type: protein

A/Residues: 1-1036 <HIG>

A/Experimental source: skin

A/Note: this is the latest in a series of papers from these workers elucidating the sequence

R/Eyre, D.R.; Glimcher, M.J.

Biochem. Biophys. Res. Commun. 48, 720-726, 1972

A/Title: Evidence for a previously undetected sequence at the carboxyterminus of the alpha1

A/Reference number: A90181; PMID:72243016; PMID:5047697

A/Accession: A90181

A/Molecule type: protein

A/Residues: 1037-1042 <EVR>

A/Experimental source: skin

A/Note: residues 1037-1042 above correspond to the carboxyl end of the protein

C/Comment: Lysines at positions 103, 700, 934, and 946 above may be hydroxylated in some

C/Comment: Most of the prolines at the third position of the tripeptide repeating unit (C

C/Comment: Pro-1002 is the only 3-hydroxyproline and the only hydroxylated proline in pos

C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

C/Keywords: coiled coil; extracellular matrix; glycoprotein; pyrogutamic acid; trimer; t

F./Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Alignment Scores:

Pred. No.:	0.00455	Length:	1042
Score:	183.00	Matches:	106
Percent Similarity:	36.34%	Conservative:	23
Best Local Similarity:	29.86%	Mismatches:	133
Query Match:	10.66%	Indels:	94
DB:	1	Gaps:	21

US-10-030-271-1 (1-909) x CGCH1S (1-1042)

```

QY 12 CCGGTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 71
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 583 ArgGlyAspProGlyPro--LysGlyAlaAspGlyAlaProGly-----LysAspGly 599
QY 72 GTGCTTCAACCGATCTTCAGAGT-----GGTGGG 101
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 600 -----LeuArgGlyLeuThrGlyProIleGlyProProGlyProIaGly 614
QY 102 CCGGCACTGACCGAGTGCAGCTGAGCTCTGCGCTTCTGCTGATGAGCTCCTGG 161
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 615 AlaProGlyAspArgGlyAlaGlyProProGlyProIaGlyProThrGlyAla--- 633
QY 162 CCGCGCGGAGGCTTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 212
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 634 -----ArgGlyAlaProGlyAspArgGlyGlyProProGlyProProGlyProIaGly 650
QY 213 ---GTTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCACT--- 266
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 651 PheAlaGlyProProGlyAla-----AspGlyGlnProGlyAlaIaIaIaIaIaIaIa 668
QY 267 -----CCTGCGGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 314
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 669 AspAlaGlyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 688
QY 315 GCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 373

```

```

Db      689 G1YProAlaGly***:::-----ValG1YAlaProG1YProG1YAlaArg 703
QY      374 GGACAGAGGAGTGGCGCGCGCGCGAGTCAAGACAGTTGCAAAATTCAGCAG 433
Db      704 G1YSerAlaG1YProProG1YAlaThrG1Y-----PheProG1Y 716
QY      434 GTCAAGGAGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGTCGGGCGCGCCA 493
Db      717 AlaAlaG1YArgValG1YProProG1YProSerG1YAsn11LeuProG1YProPro 736
QY      494 GTGGTGGTCCAGACAGCGCGCGAGAGGGCC----- 526
Db      737 G1YProAlaG1Ylys**G1YSer1YsG1YProArg1YG1YIuThrG1YProAlaG1YArg 756
QY      527 -----CAGCGGCAACCCGAGAGCAGTCAAGAGCGGCA----- 559
Db      757 ProG1YG1YProG1YProAlaG1YProProG1YProProG1YG1YIuYsG1YSerProG1Y 776
QY      560 ---GACCTTCCTCAAGAGCAAAAGTGAACCTGTCATCCGCG---TCCGGGTTCCAGCAG 613
Db      777 AlaAspG1YPro11LeuG1YAla-----ProG1YThrProG1YProG1YI11eAlaG1Y 794
QY      614 AGTACTGCGAGACAGGCGGCGACCTTGA-----GCAGGCGGTGGCATCCGCGCGGCC 667
Db      795 G1YArgG1YValG1YG1YLeuPro-G1YG1YArgG1YG1YIuArgG1YPheProG1YLeuPr 814
QY      668 AGGCGCTGGCGCGGAGAGTGAACGTTGG-----CGAGGCAACCGCAGT-- 713
Db      814 OG1YProSerG1YG1YProG1YIuYsG1YIuG1YProSerG1YAlaSerG1YIuArgG1YPr 834
QY      714 ---GCTGCGCTCAAGAGGAGCTGCGGCTGTGTGTGATCATCAAGTTCAGAGCTCT 769
Db      834 oProG1YProMetG1YProProG1YLeuAlaG1YProProG1YG1YAlaG1YArgG1YIu 854
QY      770 CCAATCTGAGACCTCTTGGGAGCAGTACCTGAGTGGCGCCT-----GCTGC 817
Db      854 YAlaProG1Y-----AlaG1YIuYAlaProG1YArgAspG1YAlaAl 868
QY      818 AGGCGCTGGCGGCGCTGTCTCTGACTGAGCGCTTGGCGGAGAGCTGTGG----- 866
Db      868 aG1YProYsG1YAspArgG1YIuThrG1YProAlaG1YProProG1YAlaProG1YAl 888
QY      867 -----CCGAGAGGCTGTTCGCTGCTGACTGAGCGCTGTCAGTGTGA 899
Db      888 aProG1YAlaProG1YProValG1YProAlaG1YIuYsAsnG1Y 902

RESULT 7
T33110
hypothetical protein C18H7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33110
R:Tin-Mollem, A.; Fronick, W.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C18H7.
A:Reference number: Z21284
A:Accession: T33110
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-460 <TIN>
A:Cross-references: UNIPROT:O9GZF7; EMBL:AF067607; PIDN:AAC17641.1; GSPDB:GN00022; CESP:
C:Genetic8:
A:Gene: CESP:C18H7.3
A:Map position: 4
A:Introns: 84/1
C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Alignment Scores:
Pred. No.: 0.000854 Length: 460
Score: 179.50 Matches: 93
Percent Similarity: 33.87% Conservative: 13

```

```

Best Local Similarity: 29.71% Mismatches: 108
Query Match: 10.46% Indels: 101
DB: 2 Gaps: 15

US-10-030-271-1 (1-909) x T33110 (1-460)
QY      36 GGAGAGAGATGATGAGCTTGAAGTCAAGAGAGTGTGCTTACCGTATGTTGAGGT 95
Db      116 G1YG1YAsnG1YMetProG1Y-----ValAlaG1YMetAlaMetAspAla 130
QY      96 GGTGGGCGGCACTGACCGAGTGCAGTGTGAGCTCTGGCCCTTCTCTGAGTGAAGC 155
Db      131 Asn11YAlaCys11LeuYsCysProAlaG1YIuProG1Y----- 144
QY      156 TCTGGGCGCGCGGAGGCTTAGCGCGCGCGGAGCGGCGCTTAGCTCTGCTGA--- 212
Db      145 -----ProProG1YProG1YIuG1YProAlaG1YProAlaG1YPro 157
QY      213 -----GCTGAGCGCGCGCGGAGTGCAGGAGAGCAACTGCGCTGAGGCA 263
Db      158 AlaG1YAsnAlaG1YAlaAspG1YG1YAlaG1YIuG1YProAlaG1YG1YAlaG1Y 177
QY      264 ACTCTGCGGCTGCTGCGCGCGGAGTGTGCTGCTGCGCGGCA----- 313
Db      178 ProPro-----G1YProProG1YProAsp1YsProG1YAlaAlaG1YSerAsp 194
QY      314 -----AGCGCGCG---GGCAGTGTCTCCAGAAAGCTATAGCTATGAGCACTCCAGCT 364
Db      195 G1YG1YAlaG1YAlaProG1YThr1YThr1YThrAsnAlaProG1YIuProG1YProAla 214
QY      365 CTTCAAGAGAGACAGAGGAGTGTGCGCGCGCGCTGCGAGTCAAGAGTTTCAATT 424
Db      214 ----- 214
QY      425 CTGACAGAGGTCACTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGTGGG 484
Db      215 -----G1YProProG1YProProG1YProAsnG1YIuAlaG1YAlaG1Y 229
QY      485 GCGGCGCGAGTGTGTGCCAGACGCGGCGGAGAGGCGCGGAGCGGAGCGGAGCGGAG--- 540
Db      230 Ser1YPro-----G1YProG1YIuG1Y--ProAlaG1YProProG1YPr 243
QY      541 -----CAGCACTCAGAGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 592
Db      243 roAsnG1YIuYsAspG1YG1YAlaG1YAlaProG1YIuYsAspG1YAspAlaG1YAlaAspG 263
QY      593 TCCGCTCCGGGTTTCAGACAGTACTGCGA-----GATGGCGGAGCTTGGAG 642
Db      263 1YAlaProG1YThrAspAlaAla1YrCysProCysProProArgSerAlaAlaLeuG1Y 283
QY      643 CAGGCGCGTGGCATCCG-----GCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGT 693
Db      283 1aG1YG1YG1YAlaG1YIuProAlaG1YAlaAspG1YAlaAlaAlaAlaAlaAlaProG1Y 303
QY      694 TTTGGGAGGCGGAGGAGTGTGCTGCTCAAGGAGGAGCGGAGCGGAGCGGAGCGGAGCGT 753
Db      303 1a---AlaProG1YAlaAlaAlaProAlaAlaG1YAlaG1Y----- 315
QY      754 AAGTTTCAGAGCTTCTCTATCTGAGACCTTCTGGGCGGAGCTGAGTGTGAGCGCGCTG 813
Db      316 -----G1YG1YAlaG1YIuProAlaG1YAlaAla 325
QY      814 CTGAGAGCGCTGCGGCGGCGGCTTCTGACTGAGCGCT-----CGGA 855
Db      325 1aProAspAlaAlaAlaAlaAlaProG1YAlaAlaAlaProAlaG1YAlaAlaAlaAlaG 345
QY      856 GAGGCTGTGGCGCGGAGGCTGTGCTGCTGCTGT 890
Db      345 IuG1YAlaG1Y---G1YG1YAlaG1YIuProAlaG1Y 355

RESULT 8
T02345
hypothetical protein KIAA0324 - human (fragment)

```

C/Species: Homo sapiens (man)
 C/Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
 C/Accession: T02345
 R/Ricke, D.O.; Bruce, D.; Munde, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
 re, J.; White, S.; Tatum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
 Submitted to the EMBL Data Library, March 1998
 A/Description: Sequencing of human chromosome 16p13.3.
 A/Reference number: Z14664
 A/Accession: T02345
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1791 <RIC>
 A/Cross-references: UNIPROT:O60382; EMBL:AC004493; NID:g2936648; PIDD:AA08453.1; PID:g2
 C/Genetics:
 A/Map position: 16
 A/Intons: 1610/2; 1706/2
 A/Note: KIA0324

Alignment Scores:
 Pred. No.: 0.000735 Length: 1791
 Score: 179.00 Matches: 89
 Percent Similarity: 37.31% Conservative: 33
 Best Local Similarity: 27.22% Mismatches: 146
 Query Match: 10.93% Indels: 59
 Gaps: 11

US-10-030-271-1 (1-909) x T02345 (1-1791)

```

QY 890 ACCGACGAGGACAGAGCCGCGCCAGCTCTCGACAGGCGCTCATGAGAACAG 831
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 699 ThrAlaArgArg---GlySerArgSerSerProGluProIuProIuSerArgThrPro 717
QY 830 CCGGACGAGGCGCTGACAGCAGGCGCCACTAGTAGTCCGCCCGC----- 786
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 718 ProArgArgArgSerSerArgSer-----SerProGluLeuThrArgGlyAla 733
QY 785 -----AAGCGCTCCAGATAGAGAGAGCTCTGAGAACCTGATGTACAACACAGAG 735
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 734 ArgLeuSerArgArgSerSerArgSerSerProGluThrArgSerArgThrPro 753
QY 734 CCGGACGCTCTGAGCAGCAGCTGCGGTGCGCTGCCCA----- 696
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 754 ProArgHis---ArgArgSerProSerValSerSerProGluProIuArgGlySerArg 772
QY 695 -----AACAGCTCCAGCTGCGCGCCAGCGCGCTGAGGCGCGCGGAGT 654
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 773 SerSerArgArgArgArgSerAlaSerSerProArgThrArgThrSerArgGly 792
QY 653 GCCACGCGCTGCTCCAGAGCTGCGGCCCATGCTCGAGACTCTGTCAACCCGAGCGCG 594
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 793 ArgSerProSerProlys-----ProArgGlyLeuGlnArgSerArgSerArgValThr 810
QY 593 ATGTCACAGGTCACCTTGCTTCAGAGGAAGTCTGGCGGCGCTTGACTGCTGGGGT 534
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 811 ArgGluArgThrArgThrArgThrArgArgArgArgSerGlySer----- 825
QY 533 GCGGCTGGGCGCCCTCTCCGCGCGCTGTGACACCACTGGCGCGCGCCGACTCCGC 474
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 826 -----SerGlnSerThrSerArgArgGlnArgSerArgSerArgSerValThr 843
QY 473 CGCTGCGCGCTTGCTGGGGGGGAGCCTGTCTCCACAGCACTGCTGCTGAGAAATTGCA 414
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 844 ArgArgArgArgArgGlyGlySerGlyTyrHisSerArgSerProAlaArgGlnGluSer 863
QY 413 CTGCTTACTGCGACGCGGACAGCACTACCTCTGCTCTTTGAAGAGCTGGAGGTG 354
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 864 ArgThrSerSerArgArgArgArgGlyArgSer----- 874
QY 353 CCATAGCTATAGCGTTGTGAGACACTGGCGCGCGCTTGCGCGCAGGTGGGAGC 294
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 875 -----ArgThrProProThrSerArgGlySerArgSerArgThrSerPro 890
QY 293 AGGTGCTGGCGGCGGACAGCGGAGAGTTGCCCGGACGCGGAGAGTTGCTCGCGG 234
  
```

```

Db 891 AlaProITrPylArgSerArgSerArgAlaSerProAlaThrHisArgArgSerArgSer 910
QY 233 CACTGCGCG-----CGGCGCTCCAGCTCCAGAGAGCTCTAGCGCGCGCC 183
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 911 ArgThrProLeuLeuSerArgArgSerArgSerArgThrSerProValSerArgArg 930
QY 182 CGGCTTAAGCTTCGCGCGCGCGCCAGAGCTTATCCAGCAGAAAGCCAGAGCTCCA-- 125
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 931 ArgSerArgSerArgThrSerValThrArgArgArgSerArgSerArgAlaSer-ProVa 950
QY 124 -GCTGCGACCTGGTCACTGTTGCCCGCCACCACTCGAACATACGAGAGAGCG----- 74
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 950 lSerArgArgArgSerArgSerArgThrProProValThrArgArgSerArgSer 970
QY 73 -----ACAGCATCCCGTAGTATCCAGGACACTTCTCTCCAGCAGCGGCGCGG 21
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 970 gThrProITrThrArgArgArgSerArgSerArgThrProProValThrArgArgSe 990
QY 20 GTCGACCCCGGATAGCGGCA 2
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 990 rArgSerArgThrProPro 996
  
```

RESULT 9

Q08B3
 BHLFI protein - human herpesvirus 4 (strain B95-8)
 C/Species: human herpesvirus 4, Epstein-Barr virus
 C/Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
 C/Accession: A03742
 R/Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
 Mol. Biol. Med. 1, 21-45, 1983
 A/Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus.
 A/Reference number: A93065; MUID:85035713; PMID:6092825
 A/Accession: A03742
 A/Molecule type: DNA
 A/Residues: 1-660 <BAN>
 A/Cross-references: UNIPROT:P03181
 R/Bear, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; He
 Nature 310, 207-211, 1984
 A/Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A/Reference number: A03794; MUID:84270667; PMID:6087149
 A/Comments: annotation: protein coding region
 C/Content: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-52;
 C/Superfamily: human herpesvirus 4 BHLFI protein

Alignment Scores:
 Pred. No.: 0.00096 Length: 660
 Score: 178.00 Matches: 105
 Percent Similarity: 34.16% Conservative: 19
 Best Local Similarity: 28.93% Mismatches: 117
 Query Match: 10.37% Indels: 122
 Gaps: 19

US-10-030-271-1 (1-909) x Q08B3 (1-660)

```

QY 2 TGCGGC-----TATCCGGGTGACCCCGCGCCGCTGTGAGAGAGATG 46
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 159 TrPArgArgArgSerGlyAlaGlnArgGlyHisProProProIuAlaGlyGlnArgPro 178
QY 47 AGTGCGTGACTACTACGGGATGCTGCTTACAGGTATGTTCAAGGTGTGGGCGGC 106
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 179 SerGlyProThr-----Gly 183
QY 107 AACTGACGAGTGCAGCTGAGACTGCTGCTTGTGCT-----GATAGAG 154
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 184 GlyArgProAlaAla-ProGlyAlaProGlyThrProAlaAlaProGlyProGlyGly 203
QY 155 CTCTGCGCGCGCGGAGGCTTAAGCCCGGCGCGGACGCGCTTAAGAGCTCTGCTGAGC 214
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 203 yAlaAlaValaProSerGlyAlaThrProHisProGlnArgGlySerGlyPro-----Al 221
QY 215 TGAAGCGCGCGGCGGAGTGG-----CGAGAGCAACT----- 248
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
  
```



```

Db      221 aAePrProAlaAlaAlaArgLeuProGluArgGlnGluProArgLeuProGlnAs 241
Qy      249 -----GCGAGTCTGCTGGGCACTCTGCGCTGCTGAGCCGCGACAGCTGCTGGCAGC 304
          |||||
Db      241 pLeuAlaAlaAlaGlnArgCysProAlaGlyProProProThrArgSerGlyAlaAla-A 261
Qy      305 TGGCGCGCAAGCGCGCGCGCGAGTCTCCAGAACGCTATAGCTATGCGACCTCCAGCT 364
          |||||
Db      261 laGlnArgThrNlaArgArgProGlyCysProArg----- 273
Qy      365 CTTCAAGAGACAGAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 418
          |||||
Db      274 --SerAlaArgAsnProGlyCysProArgThrTrpArgArgSerGlyAlaGlnArg 293
Qy      419 CAATTTCTCAGCAGGAGTACGTGAGAGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
          |||||
Db      293 lYnHsPrProProProGly-AlaGly---GlnArgProSerGlyProThrGlyAlaArgPro 311
Qy      470 ---AGCGCGAGAGTGGGGCGCGCGCGAGTGTGTCAGACGCGCGCGA----- 517
          |||||
Db      312 AlaAlaProGlyAlaProGlyThrProAlaAlaProGlyProGlyGlyAlaAlaVal 331
Qy      518 ---GAGGGCGCGCGCGCGCGCGCGAGCAGCAGTACAGAGCTGCTGCTGCTGCTGCTGCTGCTG 574
          |||||
Db      332 ProSerGlyAlaThrProAlaProGluArgGlySerGlyProAlaAspProProAlaAla 351
Qy      575 GCAAG----- 580
          |||||
Db      352 AlaArgLeuProProGluArgGlnGluProArgLeuProGlnAlaAlaAlaGln 371
Qy      581 -----TGACCTGTGACATCGGCTCGGCTCGGCTCGAGCAGAGT---ACT 619
          |||||
Db      372 ArgCysProAlaGlyProProProThrArgSerGlyAlaAlaAlaGlnArgThrNlaArg 391
Qy      620 GCGAGCAGTGGCGCGCGCTGAGAGAGCGCGCTGAGTCCCGCGCGCGCGCGCGCTG--- 676
          |||||
Db      392 ArgProProGlyCysProArgSer---AlaArgAsnProGlyCysProArgThrTrpArg 410
Qy      677 CGCGCGCAGTGAAGTGTGTCGAGAGCGCGAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 736
          |||||
Db      411 ArgArgSerGlyAlaGlnArgGlyNlaProProProGlyAlaGln----- 426
Qy      737 CTGTGTTGTGACATGACATGCTCAGAGCTCTCTATCTGAGACGCTTCTGCGGCGACT 796
          |||||
Db      427 -----ArgProSerGlyProThr 432
Qy      797 -----ACCTGAGTGGCGCT----- 812
          |||||
Db      433 GlyGlyArgProAlaAlaProGlyAlaProGlyThrProAlaAla-ProGlyProGlyGly 452
Qy      813 -----GCTGCGAGCGCTGCGGCGCGTGTCTGACTGAGCGCGCTGCGAGCGCTGCGGCGC 868
          |||||
Db      452 yGlyAlaAlaValProSerGlyAlaThrProNla-----ProGluArgGlySerGlyPr 470
Qy      869 G 869
          ||
Db      470 o 470

```

RESULT 10

F75518 hypothetical protein - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: F75518

R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Shen, M.; Vamshayavan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; MUID:20036896; PMID:10567266

A/Accession: F75518

A/Status: preliminary

A/Molecule type: DNA

```

A/Residues: 1-839 <WHI>
A/Cross-references: UNIPROT:Q9RX57; GB:AE001904; GB:AE000513; NID:G6458129; PIDN:AAF1003
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR0458
A/Map position: 1
C/Superfamily: herpesvirus immediate-early protein IE110; RING finger homology

Alignment Scores:
Pred. No.: 0.00119 Length: 839
Score: 176.50 Matches: 94
Percent Similarity: 29.47% Conservative: 23
Best Local Similarity: 23.68% Mismatches: 115
Query Match: 10.78% Indels: 165
DB: 2 Gaps: 16

US-10-030-271-1 (1-909) x F75518 (1-839)
Qy      889 CAGCAGCGAGACAGCTCCCGG-----CAACAGCTCTCGAGGAGCTCAGTCAAGA 836
          |||||
Db      258 ProAlaAlaGlnArgProAlaGlyGlyAlaProSerProAlaProAlaProAlaGlnAla 277
Qy      835 ACACGCCCGCAGG----- 821
          |||||
Db      278 AsnAlaProAlaGlySerValValProGluAlaThrValProGluSerThrProAla 297
Qy      821 ----- 821
          |||||
Db      298 AlaProSerAlaGlnThrProProThrProThrArgGluThrAlaGlnThrGluAlaSer 317
Qy      820 CTTGACAGCGCGCGCGCAGTCAAGTATGCGCGCGCGAGAG----- 782
          |||||
Db      318 ProAlaAlaProAlaAsnSerSerAlaAlaAlaProAlaGlnThrProAlaSerGluProValAla 337
Qy      781 ---GCTCAATAGAGAGAGCTGTGAGAACTTGATGACAAACACAGAGCGCGCTGCC 725
          |||||
Db      338 GlyArgProGlyThrAlaAla-----SerSerProGluSerAlaSerPro 352
Qy      724 TTGAGCGCAGACATCGCGGTGGCGCTGCCAACAAGTCAAGTCAAGCGCGCGCGCGCTGGG 665
          |||||
Db      353 ValThrValThrProAlaGlyGlyGluThrProAspThrAlaAlaSerAlaGlyThrProSer 372
Qy      664 GCGCGCGAGATG---CCAGCGCTGTCTCAAGCGTCCAGCGTCCAGTCTGAGTCTGCTC 608
          |||||
Db      373 AlaGlyArgValThrProAlaProAlaProAlaProSerAlaSerGluGlyAlaSerAlaAlaArg 392
Qy      607 GAACCGCGAGCCGAGTCTCAAGTCACTTTCCTTCAGAGAAAGTCTGGCGGCTCTG 548
          |||||
Db      393 ThrProGlyAlaGlySerGlnThrProProAlaProAlaThrProIle-----ProAla 410
Qy      547 ACTGCTGCTGGGTGGCGGCTGGG----- 524
          |||||
Db      411 ThrProAlaGlyArgSerSerGlyGluSerAlaGlyThrAlaAlaAlaArgProAlaAla 430
Qy      524 ----- 524
          |||||
Db      431 AlaProAlaProValSerGluAspArgSerAspValSerGlyLeuProArgArgGluAsp 450
Qy      523 -----CCCTCTCCGCGCGCGCTGCGACCACTGAGCGCGCGCC 482
          |||||
Db      451 AlaProAlaGluSerSerProValAlaAlaSerProAlaArgGlyAlaSerSerAlaPro 470
Qy      481 GACTCGCGCGCTCGCGTGTGTTG-----GGGGGAGAGCTGCTCT 443
          |||||
Db      471 SerSerAlaProAlaAlaAlaValProSerArgAlaProValSerGlyGlySerValSer 490
Qy      442 -----CCACTGACCTGTGAGAAATTTGACAGAACTGTGACTGCCGAC 398
          |||||
Db      491 AlaProArgThrAlaProThrAlaProValAlaGlu----- 502
Qy      397 GCGAGCGCAGCTACCTCTGCTCTTTGAGAGCTGAGGTGCCATAGCTATAGCGTT 338
          |||||
Db      503 ---GlnGlyGluValProValSer----- 509

```

```

QY 337 CTGGAGACACTGCGCCGCCCTTGCGCCGAGTGGCGGAGCTGCTGCGGCCA 278
Db 510 -----ProSerAlaAlaAlaProAlaGlyGlyAla 519
QY 277 GCACGGCGAGGAGTT-----GCCCGAGCGCGCA-----GGTTCCTCCGC 236
Db 520 SerSerAlaAlaAlaProSerAlaProAlaAlaAlaArgGlySerGlyAlaAlaGly 539
QY 235 CGCAGTCCCGCGCGCTCCAGCTCCAGCAGAGCTTAGCGCGCTGCGGCCGCTTA 176
Db 540 GlyAlaAlaGlyGlyAlaSerAlaProAlaAlaAla----- 551
QY 175 AGCTTCGGCGCGCGCCAGAGCTTATCCAGCAAAAGCCAGAGCTCCAGCTCCGACT 116
Db 552 -----ArgProAlaGlnThrProGlyAlaSerAlaGlyGly 563
QY 115 CGGTCAATTGCCGCCACCACTCGAATACGAGAGCGACAGATCCGCTAGTAGT 56
Db 564 AlaSerGlyGly-----GlyGlnGlyValSerAlaArgProSer 576
QY 55 CCAGGACTATCTCTCTCCAGCAGCGGCGCGGCTGCACCCGATAGCG 5
Db 577 GlnGly-----GlyThrProSerGlyThrProAlaSerAlaProAlaAla 591

RESULT 11
A41182
collagen alpha 1(II) chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 13-Aug-1999
C:Accession: A41182; A44885
R:McEwen, M.; Tomar, D.; de Crombrughe, B.; Vuorio, E.
J. Biol. Chem. 266, 16862-16869, 1991
A:/Title: Mouse type II collagen gene. Complete nucleotide sequence, exon structure, and
A:/Reference number: A41182; NUID:91358489; PMID:1885613
A:/Accession: A41182
A:/Status: preliminary; not compared with conceptual translation
A:/Molecule type: DNA
A:/Residues: 1-1419 <MET>
A:/Cross-references: GB:M65161
R:Cheah, K.S.; Lau, E.T.; Au, P.K.; Tam, P.P.
Development 111, 945-953, 1991
A:/Title: Expression of the mouse alpha 1(II) collagen gene is not restricted to cartilag
A:/Reference number: A44885; MUID:91347939; PMID:1879363
A:/Accession: A44885
A:/Molecule type: DNA
A:/Residues: 1-28 <CHE>
A:/Cross-references: GB:S63190; NUID:9234368; PIDN:AA19627.1; PID:9234369
A:/Note: sequence extracted from NCBI backbone (NCBI:63190, NCBI:P:63192)
C:/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:/Keywords: alternative splicing; cold; coll; extracellular matrix; glycoprotein; trime
F:/1191-1419/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Alignment Scores:
Pred. No.: 0.00109 Length: 1419
Score: 176.50 Matches: 96
Percent Similarity: 33.94% Conservative: 15
Best Local Similarity: 29.36% Mismatches: 134
Query Match: 10.29% Indels: 83
DB: 2 Gaps: 15

US-10-030-271-1 (1-909) x A41182 (1-1419)
QY 21 CCCGCGCCGCTGAGAGAGATGATGCTGACTACTAGCGGATGCTGCTGCTCA 80
Db 762 ProGlyProAla---GlyPheAlaGlyProProGly-----AlaAspGlyGln 776
QY 81 CCGTATGTTGAGAGTGT---GGCGGGCAACTGACCGAGTGGAGCTGCTGCGC 137
Db 777 ProGlyAlaAlaGlySerGlnGlyGlnAlaGlyGlnAlaGlyAlaProGly 796
QY 138 CTTTCTGCTGAGTGAAGCTCTGCGCGCGCGGAGGCTTACCGCGCGCGCGGCT 197

```

```

Db 797 ProGlnGlyProSerGly-----AlaProGlyProGlnGlyPro 809
QY 198 -----AGAGCTCTGCTGAGCT---GGA 218
Db 810 ThrGlyValThrGlyProGlySerGlyAlaArgGlyAlaGlnGlyProProGlyAlaThrGly 829
QY 219 GCGCGCGCGGAGTGGCGGAGAGCACTCGCGCTCTGGGGCAACTCTGGCGCTGT 278
Db 830 PheProGlyAlaAlaGlyValGlyProProGlyAlaAsnGlyAsnProGlyProAla 849
QY 279 GCGCGCGAGCACTGCTGCTCGCACTCGCGCGGAGCGGCGGCGGAGTCTTCAGA 338
Db 850 GlyProProGlyProAlaGlyValAspGly-----ProGlySerGly 863
QY 339 AGCTATAGCTATGAGCACTCCAGCTTCTCAAGAGAGAGAGGTAAGTCCGCTCGCG 398
Db 864 -----ArgGlyAspSerGly-----ProProGly 871
QY 399 TCGGCACTGACAGAGTTCGCAATTCTCAGACAGGCTCAGTGGAGACAGGCTCCCCC 458
Db 872 ArgAlaGlyAspProGlyLeuGlnGlyProAlaGlyAlaProGlyGlnGlyGlnPro 891
QY 459 AAC-----CAAGCGGAGCGCGGAG 479
Db 892 GlyAspAspGlyProSerGlyLeuAspGlyProProGlyProGlnGlyLeuAlaGlyGln 911
QY 480 TCGGCGCGCGCGCGAGTGTGTGTCAGACGCGCGGAGAGGGCCCGACCGCACCCA 539
Db 912 -ArgGlyIleValIleGlyLeuProGlyGlnAlaGlyGlnAlaArgGlyPheProGlyLeuProGly 931
QY 540 GCAGCAGTCAAGAGCCCGCGAGACCTTCTCTGAAGCAA---AGTGAAGTGAATCC 595
Db 931 yProSerGlyGlnProGlyValSerGlnGlyAlaProGlyAlaSerGlyAspArgGlyProPr 951
QY 596 GCGTCCGGGTTCAGAGAGTACTGCGACATGAGCGACGCTTGAGCAGAGCGGTGCAT 655
Db 951 GlnGlyProValIleGlyProProGlyLeuThrGlyProAlaGlyGlnProGlyArgGlnGlySer 971
QY 656 CCGCGCGGCC-----CCAGCGCTGCGCGCGGAGCTGAGCTGTTGGGAGGCCA 706
Db 971 rProGlyAlaAspGlyProProGlyLysArgAspGlyAlaAlaGlyValGlyAspArgGly 991
QY 707 CCGCAATGCTGCGCTCAAGGAGCTTGGCTCTGTGTTGTGATCAAGTTCTCAGAGC 766
Db 991 yGlnThrGlyAlaLeuGlyAlaProGlyAlaProGly----- 1003
QY 767 TCTCTATCTGAGCGCTTCTGGCGGAGCTACTGAGTGAGCGCTGCT---GCAGGCC 823
Db 1004 -----ProProGlySerProGlyProAlaGlyPr 1013
QY 824 TGGCGGGCGTGTCTGACTGAGCGCTTCGAGAGGCTGTGGCGGAGGCTGTTG-- 881
Db 1013 oThrGlyLysGlnGlyAspArgGlyGlnAlaGlyAlaGlnGlyPromerGlyProSerGly 1033
QY 882 -CTGCTGCTGCTGAGTGA 899
Db 1033 yProAlaGlyAlaArgGly 1039

RESULT 12
A45344
immediate-early protein - suid herpesvirus 1 (strain Kaplan)
C:/Species: suid herpesvirus 1
C:/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:/Accession: A45344
R:Vittek, C.; Kozmlik, Z.; Paces, V.; Schlrm, S.; Schwytzer, M.
Virology 179, 365-377, 1990
A:/Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op
A:/Reference number: A45344; MUID:91021039; PMID:2171211
A:/Status: translation not shown
A:/Molecule type: DNA
A:/Residues: 1-1446 <VLC>
A:/Cross-references: UNIPROT:P33479; GB:M34651; NUID:9334070; PIDN:AAA47470.1; PID:9334071

```



```

QY      480 TCGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGGGCCCCAGCCGCCACCCCA 539
      |||||
      ::|||
      |||||
Db      980 -ArgGlyIleValGlyLeuProGlyGlnArgGlyGluArgGlyPheProGlyLeuProGly 999
QY      540 GCAGCACTAGAGCCGCCAGACCTTCTCTGMAAGCA----AGTCACTGTGACATCC 595
      |||||
      ::|||
      |||||
Db      999 yProSerGlyLeuProGlyLysGlnGlyAlaProGlyAlaSerGlyAspArgGlyPro 1019
QY      596 GGCCTCCGGTTCGACGACGACTGCGACGATGGGCCAGCCTTGAGACAGGCGCTGCAT 655
      |||||
      ::|||
      |||||
Db      1019 oGlyProValGlyProProGlyLeuThrGlyProAlaGlyGluProGlyArgGlnGlySe 1039
QY      656 CCGGCGCGCC-----CGAGCGCTGGCGCGCGACCTGACGTGTTGGGCGAGCCA 706
      |||||
      ::|||
      |||||
Db      1039 rProGlyAlaAspGlyProProGlyArgAspGlyAlaAlaGlyAlaLysGlyAspArgGly 1059
QY      707 CCGGAGTCCGCGCTGCAAGGACCTGGGCTGTGGTTTGACATCAAGTTCACAGC 766
      ::|||
      |||||
      |||||
Db      1059 yGluThrGlyAlaLeuGlyAlaProGlyAlaProGly----- 1071
QY      767 TCTCTATCTGACGCTTCTGTGGGCGACTACTGAGTGGCGCCCTGCT---GCAGGCC 823
      |||||
      ::|||
      |||||
Db      1072 -----ProProGlySerProGlyProAlaGlyPr 1081
QY      824 TCGGGCGCTGTTCTGACTGAGGCTTCGAGAGGCTGTGGCGCGGAGGCTGTCG-- 881
      |||||
      ::|||
      |||||
Db      1081 oThrGlyLysGlnGlyAspArgGlyGluAlaGlyAlaGlnGlyProMetGlyProSerGly 1101
QY      882 -CCGTGCTGTGAGTGTGA 899
      |||||
      ::|||
      |||||
Db      1101 yProAlaGlyAlaArgGly 1107

```

RESULT 14

T13049

eyelid - fruit fly (Drosophila melanogaster)

C/Species: Drosophila melanogaster

C/Date: 13-Aug-1999 #sequence revision 13-Aug-1999 #text change 09-Jul-2004

C/Accession: T13049

R/Treisman, J.E.; Juk, A.; Rubin, G.M.; Heberlein, U.

submitted to the EMBL data library, March 1998

A/Reference number: Z17592

A/Accession: T13049

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-2715 <TR>

A/Cross-references: UNIPROT:Q8IN94; EMBL:AF053091; NID:g2981220; PID:g2981221; PIDN:AACC

C/Genetics:

A:Gene: eld

A/Cross-references: Flybase:FBgn0003013

C/Function:

A/Description: could act as a transcription factor antagonistic to the Wg pathway

C/Keywords: DNA binding

Alignment Scores:

Pred. No.: 0.000978 Length: 2715

Score: 176.50 Matches: 91

Percent Similarity: 30.55% Conservative: 15

Best Local Similarity: 26.22% Mismatches: 149

Query Match: 10.78% Indels: 92

DB: 2 Gaps: 14

US-10-030-271-1 (1-909) x T13049 (1-2715)

```

QY      907 CAGCCTATCCACACTGACGACGAGGCGAACAGCT---CCGCGCCACAGACCTCTCGCA 851
      |||||
      ::|||
      |||||
Db      176 GlnProHisProGlnGlnGlnProProGlnGlnProGlyGlySerProHisArg 195
QY      850 GGGCCTCAGTCA-----GGAACAGCCCGCGGCGCTGACGACGAGCGCGCAC 803
      |||||
      ::|||
      |||||
Db      196 ProProGlnGlnArgGlyIleProGlyGlnProProGlnGlyProThrProThrLeuAsn 215
QY      802 TCA-----GGTAGTCGCCCAAGAAAGCGCTCCAGATGAGAGCTCTGAGAACTTGA 752

```

```

Db      216 SerLeuLeuGlnSerSerAsnProProProProGlnHisArgGlyAlaAsnThr--- 234
      |||||
      ::|||
      |||||
QY      751 TGTCAAAACCAAGAGCCAGGCTCCCTTGAAGGACGACACTGGCGGTGGCCCAACA 692
      ::|||
      |||||
      |||||
Db      235 ---TyrAspProGlnGln---AlaAlaAlaSerAlaAlaAlaAlaAlaGlnGln 252
QY      691 CATTACAGCTCCCGCGCGGCGCTGGAGGCGCGGAGATGCAAGCCCTGTCCAAAGCTG 632
      ::|||
      |||||
      |||||
Db      253 GlnGlnAlaGlyLysProProProGlnGlnHisGlyProProProProGln----- 269
QY      631 GCCCATGCTGCAGTACTGTCTGCAACCGGAGCGGAGATGTCACAGTCACTTGGCTT 572
      |||||
      ::|||
      |||||
Db      270 -----HisGlnPro-SerProTy 275
QY      571 CAGAGAAAGTCTGGCGGCTGTGACTGTCTGGAGTGGGCTGGGCGCTTCGCC 512
      ::|||
      |||||
      |||||
Db      275 rGlyGlnGlnGlnGly-----GlyThrAlaProProProAla 287
QY      511 GCCGTCTGGACACCACTGGGCGGCGCG-----ACTCCGCGCTGCCGCTTG 461
      |||||
      ::|||
      |||||
Db      287 gProTyrSerProGlnLeuGlyProSerGlnGlnTyArgThrProProProThrAsn 307
QY      460 TTGGGGGGAGCGCTGTCCACTGACCTGTGAGATTGAGAACTGCTGACTGACC 401
      |||||
      ::|||
      |||||
Db      307 rSerArgGlyGln--SerProTyrProProAlaHisGly---GlnAsn----- 321
QY      400 GAGCGGACGCGACGACTACCTCTGTCTCTTGAAGACGTGAGGCTCATAGC 341
      |||||
      ::|||
      |||||
Db      322 -----SerGlySerTyProSerSerProGlnGlnGlnGlnGlnGlnGln 339
QY      340 GTTCTGAGACACTGGCGCGCGCGCTGGCGCGCGGACGTGCGGACGCTGCGGG 281
      ::|||
      |||||
      |||||
Db      340 GlnGlnGln-----AlaGlyGlnGlnProGlyGlyProValProGlyGlyPro 355
QY      280 CCAAGACGCGCAGAGATGCCCCA----- 257
      |||||
      ::|||
      |||||
Db      356 ProProGlyThrGlnGlnProProGlnGlnAsnThrProProThrSerGlnTySer 375
QY      256 -----GCAGCCGACAGTGTCTCTCGCGC 233
      |||||
      ::|||
      |||||
Db      376 ProTyProGlnArgTyProThrProProGlyLeuProAlaGlyGlySerAsnHisArg 395
QY      232 ACTGCCCGCGCGCT-----CCAGCTCAGACGAGACTCTA 197
      |||||
      ::|||
      |||||
Db      396 ThrAlaTySerThrHisGlnTyProGluProAsnArgProTyProGlyGlySerSer 415
QY      196 GGCCTGCGGCGCGCGCTGAGCGCTCGGCGGCGGCGGACGAGGCTCATCCAGAGAAAG 137
      |||||
      ::|||
      |||||
Db      416 ProSerProGlySerGly-----HisProLeuProProAlaSer 428
QY      136 CAGAGACTCAGCTCGACTCGGTGATGTCGCCGCGCACACCTCGAACATACGTTGAA 77
      |||||
      ::|||
      |||||
Db      429 ProHisHisValProProLeuGlnGlnGlnProProProProProHisHisValSerAlaGly 448
QY      76 GCGACAGACATCCGTAGTACTCAAGGACTCATCTCTCCAGACAGGCGCGGGCTG 17
      |||||
      ::|||
      |||||
Db      449 GlyProProProSerSerSerProGlyHisAlaProSerProSerProGlnProSerGln 468
QY      16 ACCCGATAGCGCCA 2
      |||||
      ::|||
      |||||
Db      469 AlaSerProSerPro 473

```

RESULT 15

CGB01S

collagen alpha 1(I) chain - bovine (tentative sequence) (fragments)

C/Species: Bos primigenius taurus (cattle)

C/Date: 24-Apr-1984 #sequence revision 31-Dec-1993 #text change 09-Jul-2004

C/Accession: A91193; A91229; A91387; A91211; A91201; A91200; A43048; A02853

R/Rautenberg, J.; Timpl, R.; Furthmayr, H.

Eur. J. Biochem. 27, 231-237, 1972

A/Title: Structural characterization of N-terminal antigenic determinants in calf and hu

A/Reference number: A91193; MUID:72255334; PMID:4115172

Thu Mar 24 07:06:36 2005

us-10-030-271-1.rpt

Page 14

Search completed: March 22, 2005, 01:54:54
Job time : 51.3945 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 09:08:14 ; Search time 340.586 Seconds
(without alignment)
9046.504 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 1883
Sequence: 1 aggcgcataatagagaag.....ctggaaggcatagtg999 1883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/BACKFILE1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	719	38.2	847	4	US-09-023-655-182
2	628	33.4	1067	4	US-09-620-312D-853
3	206.2	11.0	1142	4	US-09-733-167A-2
4	204.6	10.9	1045	3	US-08-859-167-3
5	204.6	10.9	1045	3	US-09-109-273-3
6	204.6	10.9	1045	3	US-09-276-993-3
7	204.6	10.9	1045	3	US-09-723-450-3
8	201.2	10.7	1142	4	US-09-733-167A-4
9	201.2	10.7	1200	3	US-08-859-167-5
10	201.2	10.7	1200	3	US-09-109-273-5
11	201.2	10.7	1200	3	US-09-276-993-5
12	201.2	10.7	1200	4	US-09-723-450-5
13	156.4	8.3	630	4	US-09-733-167A-8
14	88.4	4.7	342	4	US-08-233-463-14
15	56.2	3.0	7218	1	US-08-233-463-14
16	50.8	2.7	1151	1	US-07-704-288C-2
17	50.8	2.7	1151	1	US-08-093-372-1
18	50.8	2.7	1151	1	US-08-379-259-2
19	49.8	2.6	999	4	US-09-902-540-3684
20	49.8	2.6	16924	4	US-09-902-540-1178
21	49	2.6	929	4	US-09-902-540-1093
22	49	2.6	15789	4	US-09-902-540-1139
23	48.2	2.6	861	4	US-09-902-540-7313
24	48.2	2.6	5764	4	US-09-902-540-693
25	47.8	2.5	50937	3	US-09-428-517-1
26	47.4	2.5	3957	4	US-10-237-551-193
27	47.4	2.5	154746	4	US-09-827-688-8

C 28	47.4	2.5	154746	4	US-09-827-688-8	Sequence 8, Appli
C 29	47	2.5	603	4	US-09-902-540-9070	Sequence 9070, Ap
C 30	47	2.5	7719	4	US-09-902-540-969	Sequence 969, Ap
C 31	46.8	2.5	774	4	US-09-266-965-43	Sequence 43, Appl
C 32	46.8	2.5	1515	4	US-09-902-540-8120	Sequence 8120, Ap
C 33	46.8	2.5	7846	4	US-09-902-540-830	Sequence 830, Ap
C 34	46.8	2.5	53500	4	US-09-266-965-76	Sequence 76, Appl
C 35	46.6	2.5	1509	4	US-09-724-797-89	Sequence 89, Appl
C 36	46.4	2.5	9937	4	US-09-902-540-946	Sequence 946, Ap
C 37	46.2	2.5	663	4	US-09-902-540-7269	Sequence 7269, Ap
C 38	46.2	2.5	4324	4	US-09-902-540-684	Sequence 684, Ap
C 39	45.8	2.4	699	4	US-09-902-540-9109	Sequence 9109, Ap
C 40	45.8	2.4	831	4	US-09-252-991A-1479	Sequence 1479, Ap
C 41	45.8	2.4	1962	4	US-09-252-991A-1693	Sequence 1693, Ap
C 42	45.8	2.4	11963	4	US-09-902-540-1072	Sequence 1072, Ap
C 43	45.8	2.4	43280	2	US-08-804-227C-1	Sequence 1, Appli
C 44	45.6	2.4	601	4	US-09-949-016-19456	Sequence 19456, A
C 45	45.6	2.4	601	4	US-09-949-016-19457	Sequence 19457, A

ALIGNMENTS

RESULT 1
US-09-023-655-182
Sequence 182, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
TITLE OF INVENTION: Jeffrey J. Seilhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 182:
SEQUENCE CHARACTERISTICS:
LENGTH: 847 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: FIBRNOT01
CLONE: 053532
US-09-023-655-182
Query Match 38.2%, Score 719, DB 4, Length 847.

Best Local Similarity 93.0%; Pred. No. 2.9e-168; Matches 788; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY	1096	GCCTCCGATCCAGAGCTGGCAGAGATTGATCCACCTCCCAAGTCTCCGGGCCACTTCTC	1155
DB	1	GCCTCCGATCCAGAGCTGGCAGAGATTGATCCACCTCCCAAGTCTCCGGGCCACTTCTC	60
QY	1156	CTGGGAGAGACGACATCTCTA-----	1176
DB	61	CTGGGAGAGACGACATCTCTA-----	1176
QY	1177	-----CCCTTGAGAGCCCTCCCAAGAGTGGGCTCTAGGC	1216
DB	121	GACAGAGCCGGGACAGAGGCTCTTGACAGCCCTCCCAAGAGTGGGCTCTAGGC	180
QY	1217	CTAAACATTTCCAGCTGAGTTCTTCCAGACTCTCCCTACCCAGGTGCCCCCT	1276
DB	181	CTAAACATTTCCAGCTGAGTTCTTCCAGACTCTCCCTACCCAGGTGCCCCCT	240
QY	1277	TAGCCTCCGAGGCGGGGCTGGGCTGTATCTCAAGAGGAGGGGACAGCTACACT	1336
DB	241	TAGCCTCCGAGGCGGGGCTGGGCTGTATCTCAAGAGGAGGGGACAGCTACACT	300
QY	1337	CACCAAGGCGCCCTGACATGTATCTGTATCTTGGGCTGTGCACTGTACAGGT	1396
DB	301	CACCAAGGCGCCCTGACATGTATCTGTATCTTGGGCTGTGCACTGTACAGGT	360
QY	1397	GACACACTGCTCATAGTCTCACATGCCCCCTGCTGAGATTTCCCTGGGCTTGCCTG	1456
DB	361	GACACACTGCTCATAGTCTCACATGCCCCCTGCTGAGATTTCCCTGGGCTTGCCTG	420
QY	1457	GCTCTGCTCCGACACACTTCTTGGCTTGAAGGCTTCTCTGACAGCTCTAATT	1516
DB	421	GCTCTGCTCCGACACACTTCTTGGCTTGAAGGCTTCTCTGACAGCTCTAATT	480
QY	1517	GACCAACCAACTGGGCTTTCAGCATGAGTGGGCACTGGAGTGGGCTGACATG	1576
DB	481	GACCAACCAACTGGGCTTTCAGCATGAGTGGGCACTGGAGTGGGCTGACATG	540
QY	1577	GAGCTGCTCACTTGGCCACATCTCCAGCCAGGAGGCTGAGCTTCAATT	1636
DB	541	GAGCTGCTCACTTGGCCACATCTCCAGCCAGGAGGCTGAGCTTCAATT	600
QY	1637	CAGACTGACTCTCTCACTTCCCTGCTGTCTCAAGCTGACATGACTTGA	1696
DB	601	CAGACTGACTCTCTCACTTCCCTGCTGTCTCAAGCTGACATGACTTGA	660
QY	1697	GATATGCACTGAGTGTCAATGGAGTGTATGGAGATGATACCAAGGCTTCTGT	1756
DB	661	GATATGCACTGAGTGTCAATGGAGTGTATGGAGATGATACCAAGGCTTCTGT	720
QY	1757	TGCAATGGGGGCAAAACAGTAACAGCACCCTTCTGGAAAGGATGCAAGGCTT	1816
DB	721	TGCAATGGGGGCAAAACAGTAACAGCACCCTTCTGGAAAGGATGCAAGGCTT	780
QY	1817	GGGGGATGATGAAAGACTTTTCAATGATACCAATTAAGCTGCTGGAAAGGCAAT	1876
DB	781	GGGGGATGATGAAAGACTTTTCAATGATACCAATTAAGCTGCTGGAAAGGCAAT	840
QY	1877	AGGTGGG 1883	
DB	841	AGGTGGG 847	

RESULT 2

US-09-620-312D-853
 ; Sequence 853, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wehrman, Tom
 APPLICANT: Xue, Aidong J.
 APPLICANT: Yang, Yonghong
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yungqing
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhilwei
 APPLICANT: John Tillinghast
 APPLICANT: Drmanac, Radoje T.
 TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/09/620,312D
 PRIOR FILING DATE: 2000-07-19
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1105
 SOFTWARE: pt_FL_genes Version 1.0
 SEQ ID NO 853
 LENGTH: 1067
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (151)..(984)
 US-09-620-312D-853

Query Match 33.4%; Score 628; DB 4; Length 1067;
 Best Local Similarity 96.3%; Pred. No. 1.1e-145;
 Matches 659; Conservative 0; Mismatches 10; Indels 15; Gaps 1;

QY	37	AACAAGTGGCTCGGCGCTGCGCCAGAGTCAATCGAGCCGCAATCTGGCGGGTTCTG	96
DB	64	AACAAGTGGCTCGGCGCTGCGCCAGAGTCAATCGAGCCGCAATCTGGCGGGTTCTG	123
QY	97	AGCTGTTCCGCTCCCTCCCGGGGAATGGGCTATCCGGGTGACCCCGGCGCGTGC	156
DB	124	AGCTGTTCCGCTCCCTCCCGGGGAATGGGCTATCCGGGTGACCCCGGCGCGTGC	183
QY	157	TGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGCTTACCGTATGTTGAG	216
DB	184	TGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGCTTACCGTATGTTGAG	243
QY	217	TGGTGGGCGGGCAACTGACCGAGTGGAGCTTCTGGCTTCTGCTGATGAG	276
DB	244	TGGTGGGCGGGCAACTGACCGAGTGGAGCTTCTGGCTTCTGCTGATGAG	303
QY	277	GCTCCTGGGCGCGCGGAGGCTTACCGCGGCGCGGCGGCTAGAGCTCTGCTGAG	336
DB	304	GCTCCTGGGCGCGCGGAGGCTTATCCGGGCGCGGCGGCTTAAAGCTCTGCTGAG	363
QY	337	CTGAGCGCGCGGCGGAGTGGCGGAGAGCAACTGCGGCTGCTGGGCAACTCTGCGC	396
DB	364	CTGAGCGCGCGGCGGAGTGGCGGAGAGCAACTGCGGCTGCTGGGCAACTCTGCGC	423
QY	397	GTGCTGGCGCGCAGCACTGCTGCGGCACTGGCGCGGCGGCGGCGGCAAGTGTCT	456
DB	424	GTGCTGGCGCGCAGCACTGCTGCGGCACTGGCGCGGCGGCGGCGGCGGCAAGTGTCT	483
QY	457	CCAGAACGTATAGCTATGAGCACTCCAGCTCTTAAAGAGAGAGAGAGGTGAGCGCT	516
DB	484	CCAGAACGTATAGCTATGAGCACTCCAGCTCTTAAAGAGAGAGAGAGGTGAGCGCT	543
QY	517	CGCGGTGGAGTCAAGCAAGTCTGCAAAATCTGACAGAGGTCAAGTGGAGAGAGGCTCC	576
DB	544	CGCGGTGGAGTCAAGCAAGTCTGCAAAATCTGCAAAATCTGACAGG-----GGCTCC	588

QY 577 CCCCCAACGAGCGGCGGAGTCCGGGCGGCCCACTGTGTGCCAGACGGCGG 636
DB 589 CCCCCAACGAGCGGCGGAGTCCGGGCGGCCCACTGTGTGCCAGACGGCGG 648
QY 637 CGGAGAGAGGCGGCGGAGTCCGGGCGGCCCACTGTGTGCCAGACGGCGG 696
DB 649 CGGAGAGAGGCGGCGGAGTCCGGGCGGCCCACTGTGTGCCAGACGGCGG 708
QY 697 GGCMAAGTGAAGTGTGACATCCGG 720
DB 709 GGCMAAGTGAAGTGTGACATCCGG 732

RESULT 3
US-09-733-167A-2
Sequence 2, Application US/09733167A
Patent No. 6696547
GENERAL INFORMATION:
APPLICANT: Peter, Marcus
TITLE OF INVENTION: Protein for Regulation of Apoptosis
FILE REFERENCE: 4121-120
CURRENT APPLICATION NUMBER: US/09/733,167A
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: PCT/DE99/01712
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
PRIOR FILING DATE: 1998-06-08
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1142
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1042)..(1042)
OTHER INFORMATION: n represents any one of a, c, t, and g.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1043)..(1043)
OTHER INFORMATION: n represents any one of a, c, t, and g.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1114)..(1114)
OTHER INFORMATION: n represents any one of a, c, t, and g.
US-09-733-167A-2

Query Match 11.0%; Score 206.2; DB 4; Length 1142;
Best Local Similarity 52.4%; Pred. No. 3.2e-41;
Matches 504; Conservative 0; Mismatches 448; Indels 9; Gaps 2;

QY 103 TTCGCTCCCTCCCGGGAATGCGCTATCCGGTGCACCCCGCGCTGCTGGAG 162
DB 7 TTGGTCTCTGAGCGCCCGAGATGCGGCTTAAGCGCGCGCAAGCAGTGTGCCA 66
QY 163 GAGGATAGGCTGAGCTACTAGGATGCTGTGCTTACCGTATGAGTGTG 222
DB 67 GAGGATAGGCTGAGCTACTAGGATGCTGTGCTTACCGTATGAGTGTG 126
QY 223 GCGGCGGAGTACCGGAGTGCAGTGCCTGCTGCTTCTGTGATGAGTCTCT 282
DB 127 GCGGCGGAGTACCGGAGTGCAGTGCCTGCTGCTTCTGTGATGAGTCTCT 186
QY 283 GCGGCGGAGTACCGGAGTGCAGTGCCTGCTGCTTCTGTGATGAGTCTCT 342
DB 187 GATGACCGAGGCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 243
QY 343 GCGGCGGAGTACCGGAGTGCAGTGCCTGCTGCTTCTGTGATGAGTCTCT 402
DB 244 GCGGCGGAGTACCGGAGTGCAGTGCCTGCTGCTTCTGTGATGAGTCTCT 303
QY 403 GCGGCGGAGTACCGGAGTGCAGTGCCTGCTGCTTCTGTGATGAGTCTCT 462

DB 304 ACTGCGCACGACCTGCTGCTTACGTCACCTCAAGAGAGAGAGGAGTGTGCTCCCTGAT 363
QY 463 CGCTATGCTATGAGCACTTCACGCTTCAAGAGAGAGAGTGTGCTCCCTGAT 522
DB 364 CTGTGAGCAAGTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
QY 523 CGGAGTGAAGAGAGTCTGCAAAATTCAGCAGAGAGAGAGAGAGAGAGAGAG 582
DB 424 GATGACAG 477
QY 583 ACCAAGCGAG 642
DB 478 GTGCTGTGTGCTCCCTGCTGCTTCAAGAGAGAGAGAGAGAGAGAGAGAG 537
QY 643 GCGGCGGAG 702
DB 538 GCGGCGGAG 597
QY 703 GTGAGCTGTGAG 762
DB 598 GAG 657
QY 763 GAG 822
DB 658 GAG 717
QY 823 GAG 882
DB 718 GAG 777
QY 883 TCAAGCTGTGAG 942
DB 778 TCAAGCTGTGAG 837
QY 943 GCGGCGGAG 1002
DB 838 GAG 897
QY 1003 CGGCTGTGAG 1062
DB 898 AAGCTGTGAG 957
QY 1063 A 1063
DB 958 A 958

RESULT 4
US-08-859-167-3
Sequence 3, Application US/08859167
Patent No. 6037461
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,167

FILING DATE: 435
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Deluca, Mark
 REGISTRATION NUMBER: 33,229
 REFERENCE/DOCKET NUMBER: TJU-
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1045 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: both
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 88..1044
 US-08-859-167-3

Query Match 10.9%; Score 204.6; DB 3; Length 1045;
 Best Local Similarity 52.3%; Pred. No. 7,8e-41;
 Matches 503; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

QY 103 TTCGCGCTCTCCCTCCCGGGAATGGCGCTATCCGGCTGACCCCGCGCCCTGCTGGAG 162
 DB 67 TTGCGTACCTGAGACCCCGACATGGCGGCTTAAAGCGCGGCAAGCCAGGTGTGCCA 126
 QY 163 GAGGAGATGAGCTGTGACTCTACTGAGGATCTGCTGCTTACCGGTATGTTGAGGATG 222
 DB 127 GAAGAGCATGTGAGACGGAACATGGGCTGTACAGCTGACGCAATGTTGACATGAG 186
 QY 223 GCGCGCAACTGACCGAGTGCAGCTGAGACTCTGCTGCTTCTGCTGATGAGGCTCT 282
 DB 187 GGCACATCATGTGACACAGAGATGTCGGCTCTTCTTCTTGTGATGATCATT 246
 QY 283 GCGCGCGCGGAGGCTTACCGCGGCGCGAGCGGCTTACGCTCTGCTGAGCTGAG 342
 DB 247 GATGACACGAGGCTGACTCAT---CCGAATGACGTGACTTATTTGCGACTGAG 303
 QY 343 CGCGCGGCGGAGTGGCGGAGACACTGCGGCTGCTGGGCAATCCCGCGGCTG 402
 DB 304 CGCAGGCGCGCTGTATGAAAGTAACTTTCGCAAGTGTGAGCTGCTGCGCATCATC 363
 QY 403 GCGCGCACGACTGTGCTGCGCACTGCGCGAGCGCGCGCGCACTGTCTCGAGA 462
 DB 364 ACTGCGCAGGACCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 423
 QY 463 CGCTATGCTATGAGCACTTCAGCTCTTCAAAGAGAGAGAGAGAGAGAGAGAGAG 522
 DB 424 CTGTGATGACAAAGTATCTGAGAGAGACATCAATTCGCTATGTGACCCCGAGGCTCTCAGT 483
 QY 523 CGCGAGTCAAGCACTTCTGCAAAATTCACAGAGGTGAGGAGAGAGAGAGAGAGAG 582
 DB 484 GATTCAGAACCAAGGCC---TCCCGAGCCCTCTTAAACAGTGTCTCCCACTATCTT 537
 QY 583 ACCAAGCGCAGCGCGGAGTTCGCGGCGCGCGCAAGTGTGTGCTGAGAGAGAGAG 642
 DB 538 GTGTGTGTGTGCGCCACTTGGGCTCTCAATGTGTAGCAAGGCGCGCGAGAGAGAG 597
 QY 643 GGGGCGCGCGCGAGCGCGAGCACTCAAGAGCGCGCGAGAGAGAGAGAGAGAGAG 702
 DB 598 GCGACACTTGGGAGCGAGCAAAACGCGGAGGTGACAGATCCAGATCCAGAGAGAG 657
 QY 703 GTGACCTGTGACATTCGCGGCTCGGCTTCGAGCAGAGTACTGAGAGAGAGAGAGAG 762
 DB 658 CAGACATGTATCAATCAAGTGTGCGGCTTGGGCTGAAATCTGCAAGCAGTGAAGTCTG 717
 QY 763 GAGCAGGCGCTGTGATTCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822
 DB 718 CAGGCGCATGTCTTCTCTAACAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777

QY 823 CAGGCGCAGCGAGTGTGCTGCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 882
 DB 778 CAGGCGCAGCAGCAGTCTTCAAGTCCCGGAGCTGGGCTCATCATCTGTGACATCAAGTTC 837
 QY 883 TCAGAGCTCTCTATCTGAGAGCGCTTCTGGGCGAGTACTAGTGGCGCGCTGTGAG 942
 DB 838 TCTGAGCTACCTTACTGATGATGATTTCTGCGCTGATCATCAATGAGCTTTATTAGAG 897
 QY 943 GCGCTGCGGCGGCTGCTTCTGACTGAGAGCGCTGCGAGAGAGCTGTGAGCGAGAGCTGTT 1002
 DB 898 GCACCTTAAAGTGTCTTCAACAGACTCCCTCAAGCAAGCTGTGGCGCATGAGCATTC 957
 QY 1003 CGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1062
 DB 958 AAGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
 QY 1063 A 1063
 DB 1018 A 1018

RESULT 5
 US-09-109-273-3
 Sequence 3, Application US/09109273
 Patent No. 6063760

GENERAL INFORMATION:

APPLICANT: Alnemri, Emad S.

TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF

TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760ris

STREET: One Liberty Place, 46th floor

CITY: Philadelphia

STATE: PA

COUNTRY: USA

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: WINDOWS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/109,273

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/859,167

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: TJU-

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1045 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 88..1044

US-09-109-273-3

Query Match 10.9%; Score 204.6; DB 3; Length 1045;
 Best Local Similarity 52.3%; Pred. No. 7,8e-41;
 Matches 503; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

```

QY 103 TTCGCTCCCTCCCGGGAATGCGCTATCCGGGTGACCCCGGCGCGTGTGGAG 162
Db 67 TTGGTACTGAGCGCCCGAGATGGCGGGCTAAAGGGGGGGAAGCGAGGTGGCCA 126
QY 163 GAGGATAGTGGCTGAGCTACTACGGAATGCTGCTTCAACGTAATGTTGAGGTGTG 222
Db 127 GAAGAGATGATGAGCAGGAACAATGGGCTGTACAGCCCTGACCCCAATGTTGACATG 186
QY 223 GGGGGCACTGACCGAGTGGAGCTGAGAGCTCTGGCTTTTGTGATGAGGCTCT 282
Db 187 GGGCACTATGTGACACAGAGATGTCGCTGCTTTCTTCTTGTGTGATGTCATT 246
QY 283 GGGCGCCGCGAGGCTTAAAGCCGCGGAGCGGCTTAAAGCTCTGCTGAGCTGAG 342
Db 247 GATGACCAAGAGCGGTGACTAT---CCGAATGAGAGTGACTTTATTGGCACTGGAG 303
QY 343 GCGCGCGGAGTGGCGGAGAGCAACCTGCGCTGCTGGGGCAACTCTGCGGTGCTG 402
Db 304 GCGCAGGGCGGCTGTGATGAAGTAACTTTGCGCAGGTGTGCACTGCTGCGCATATC 363
QY 403 GCGCGGAGCACTGCTGCGGCACTGCGGAGAGGCGGCGGCGGCAAGTCTCCAGAA 462
Db 364 ACTGCGCAGCACTGCTGCGCTTACGTCACTCAAGAGAGAGCGGCGCTGTGTGCTGAT 423
QY 463 CGCTATAGTATGAGCACTTCCAGCTTTTCAAAAGAGACAGAGGAGTACTGCGCGCT 522
Db 424 CTGTGATACAGATATCTGAGAGAGACATTCATTTGCTATGATGAGACCCCAAGCCCT 483
QY 523 GCGGAGTCAAGCACTTGTGCAAAATTTGAGAGGAGTCAAGTGGAGAGACAGCTCC 582
Db 484 GATTCAGAACCAAGGCGC-----TCCCGAGCCCTTAAACAGTGCCTCCCACTATCT 537
QY 583 ACCAAGGGGAGCGGGGAGTCCGGGCGCGCCAGGTGTGTCACACCGCGGCGGAGA 642
Db 538 GTGTGTGTGTCGCCACTTGCGGGTCTCAAGTGTGACAGCGGCGAGCGCGGAGAGA 597
QY 643 GGGGCGCCAGCGGAGCCGAGCAGACAGTCAAGGCGCGGAGCTTCTCTGAAGGAGAA 702
Db 598 GCGCACAATTGGAGCCAGCCAGAAACGCGGAGTCAATGACACAGATCCCAAGAGAGAG 657
QY 703 GTGACTGTGACATCCGCGCTCCGGTTTGAAGCAGAGTACTGCGAGCATGGGCGCCT 762
Db 658 CAGACATGTGACATCAGATGCGGGTTTGGGCTGAATACTGCCAGCATGAGACTGCT 717
QY 763 GAGCAGGCGGTGGATCCCGCGGCGCCAGGCGGTGCGGCGGCGCTGAGAGTGTGGG 822
Db 718 CAGGGCAATGTTCTTAAACAGAGAGACCCACTTGAAGCGCCAGTTTGAAGCGCTTAA 777
QY 823 CAGGCGCAGGAGTGTGCGCTCAAGGAGACCTGGGCTCTGTGTTGTGACATCAAGTTC 882
Db 778 CAGGCGCAACATCTCACTCAAGTCCCGGAGCTGGGCTTCACTGTGACATCAAGTTC 837
QY 883 TCAGAGCTCTCTATCTGAGCGCTTGTGGGCGACTACCTGAGTGGCGGCTGTGAG 942
Db 838 TGTGAGCTCACTCACTGATGATCTGTGGGTGACTACATCAATGAGCTTTTATTGAG 897
QY 943 GCGCTGGGGCGGCTTCTGTGACTGAGGCGCTGTGAGAGGCTGTGGGCGGAGGCTGT 1002
Db 898 GCACTTAAAGTGTCTTATCAACAGACTCCCTCAAGAGAGTGTGGGCGCATGAAGCATC 957
QY 1003 GCGCTGTGTGCACTGTGATGAGAGCTGATATGAGGCTGGCGGCGCGCTTGTGCTG 1062
Db 958 AAGTGTCTGTAAATGTAGAGAGAGAGACTATGAGCTGGGCGCAGAAATCTCTGAGG 1017
QY 1063 A 1063
Db 1018 A 1018

```

RESULT 6
US-09-276-993-3
; Sequence 3, Application US/09276993

```

; Patent No. 6207801
; GENERAL INFORMATION:
; APPLICANT: Alnemir, Emad S.
; APPLICANT: Fernandez-Alnemir, Teresa
; TITLE OF INVENTION: RAD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6207801r1s
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,993
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/859,167
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1045 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..1044
; US-09-276-993-3
;
Query Match 10.9%; Score 204.6; DB 3; Length 1045;
Best Local Similarity 52.3%; Pred. No. 7.8e-41;
Matches 503; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

```

```
QY 463 CGCTATAGCTATGGCACTCCAGCTCTTCAAGAGACAGAGGGTAGTGGCTGGCGCT 522
DB 424 CTGTGATGACAAATATCTGGAGAGACATCAATTCGTATGTGACCCCGAGAGCCCTCAGT 483
QY 523 CGGCACTCAAGCAAGTCTGCAATTCACGAGGAGTCAAGTGGAGACAGGCTCCCCCA 582
DB 484 GATCCAGAACCAAGGCC-----TCCCCAGCCCTTAAACAGTGCCTCCCACTATCCCT 537
QY 583 ACCAAGCGGACGAGCGGAGTCCGGGCGCGCCAGTGTGTGTCCAGACGCGCGCGAGA 642
DB 538 GTGTGTGTGTGCCCACTTCGGGTCTCTCAGATGTGTAGCAAGCGGCGACGCGAGGAGA 597
QY 643 GGGGCCCCGAGCGGACCCCAAGCAGTCAAGAGCCCGGCAACCTTCTTGAAGCAAA 702
DB 598 GCCACACTTGGAGCGGACGAGAAACCGCGAAGTCAGTGACACAGATCCCAAGAGAAAG 657
QY 703 GTGACCTGTGACATCCGGGCTCGGGGTTCCAGCAGAGTCACTGGAGGATGGGCGGAGCTTG 762
DB 658 CAGACATGTGACATCAAGACTGCGGGGTTCCGGGCTGAAATCTGCCAGCATGAGACTGCTCTG 717
QY 763 GAGCAGGCGCTGCGATCCCGGCGCGCCCAAGCGCTGCGCGGAGCTGACGTGTTGGG 822
DB 718 CAGGGCAATGTCTTCTTAACAAGCAGGAGCCACTTGAAGGCCAGTTTGAAGCCTTTAAC 777
QY 823 CAGGGCAACGCACTGCTGCGCTCAAGGAGCTGGGCTCTGTGTGTGTGACATCAAGTTT 882
DB 778 CAGGGCAACACACTCTCAAGTCCCGGAGCTGGGCTCCATCATGTGTGACATCAAGTTT 837
QY 883 TCAGAGCTCTCCATCTGAGCGGCTTCTGGGGGAGCATCACTGATGGCGCCCTGCTGAG 942
DB 838 TCTGAGCTACCTCACTCGATCGATGATTTCTGCGTGAATCAATCAATGAGCTTTATTAGAG 897
QY 943 GCCCTGCGGGGCTGTCTTCTGACTGAGGCCCTGCGAGAGGCTGTGGGCCGAGAGGCTGTT 1002
DB 898 GCATTTAAAGGTGTCTTCAATCAAGACTCCCTCAAGAGCTGTGGGCCATGAGGCATC 957
QY 1003 CGCCTGTGCTCAGTGTGATGAGGCTGACTATGAGGCTGGCGGCGCCCTGTTGCTG 1062
DB 958 AAGCTGTGTAATGTAGAGAGGAGACTATGAGTGGGCCGAGCAAAATCTCTGAGG 1017
QY 1063 A 1063
DB 1018 A 1018
```

RESULT 7

US-09-723-450-3
Sequence 3, Application US/09723450

Patent No. 6576751
GENERAL INFORMATION:

APPLICANT: Alnemet, Emdad S.
TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, At

FILE REFERENCE: T02445
CURRENT APPLICATION NUMBER: US/09/723,450

CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 09/276,993

PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 08/859,167

PRIOR FILING DATE: 1997-05-20
NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn version 3.0
SEQ ID NO 3

LENGTH: 1045

TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: misc_feature
OTHER INFORMATION: No. 6576751el Sequence

NAME/KEY: CDS
LOCATION: (88)..(1044)

US-09-723-450-3

```
Query Match 10.9%; Score 204.6; DB 4; Length 1045;
Best Local Similarity 52.3%; Pred. No. 7,8e-41;
Matches 503; Conservative 0; Mismatches 449; Indels 9; Gaps 2;

QY 103 TTCCTGCTCCCTCCCCCGGGAATGGCGCTATCCGGGTGACACCCGCGCCCTGTGGAG 162
DB 67 TTGCGTACTTGAAGCCCCCAGCATGGCGGCTTAAACCGCGGAGCAAGCCAGGTGTGCCA 126
QY 163 GAGATGATGTGCTGAGTCACTAATCAAGGATCTGTGCTTCAACGATGTTCGAGGTGTG 222
DB 127 GAAAGCATGTGTGAGAGAGAACATGGGCTGTACAGCTGTCAACGATGTTTACATCGTG 186
QY 223 GGGCGGCACTGACCGAGTGCAGCTGAGCTTCTGTGCTTCTGTGTGTATGAGCTCT 282
DB 187 GGCATCATCTACACACAGAGATGTGCGGCTTCTTCTTCTTGTGTATGTGAT 246
QY 283 GGGCGCGCGGAGGCTTACCCCGGCGCGAGCGGCTTGAAGCTCTGTGAGCTGAG 342
DB 247 GATGACACAGAGCTGTGATCTAT---CCGAAATGACGTGACTTCTTATTGGCATGGAG 303
QY 343 CGCGCGGCGAGTGGCGCGAGAGCAACTGCGCTGTGGGCAACTCCGTGCGGCTGTG 402
DB 304 CGCGAGGGCGGCTGTGATGAAAGTAACTTTCGCAAGTGTCTGACGCTGTGCGCATATC 363
QY 403 GCCCGCCAGCACTCTGTGCTGCACTGTGCGGCAAGCGCGCGGCACTGTCTCAGAA 462
DB 364 ACTCGGCAGACCTGTGCTGCTTACGTCACCTCAAGAGAGAGCGGCTGTGCGCTGAT 423
QY 463 CGCTATAGCTATGGCACTCCAGCTCTTCAAGAGAGACAGAGGTAGTGCCTGTGGCGCT 522
DB 424 CTGTGATGACAAATATCTGGAGAGACATCAATTCGTATGTGACCCCGAGAGCCCTCAGT 483
QY 523 CGGCACTCAAGCAAGTCTGCAATTCACGAGGCTCACTGGAGACAGGCTCCCCCA 582
DB 484 GATCCAGAACCAAGGCC-----TCCCCAGCCCTTAAACAGTGCCTCCCACTATCCCT 537
QY 583 ACCAAGCGGACGAGCGGAGTCCGGGCGCGCCAGTGTGTGTCCAGACGCGGCGAGA 642
DB 538 GTGTGTGTGTGCCCACTTCGGGTCTCTCAGATGTGTAGCAAGCGGCGACCGGAGGAGA 597
QY 643 GGGGCCCCGAGCGGACCCGAGCAGTCAAGGCCGCGAGACTTCTCTGAAAGCAAA 702
DB 598 GCCACACTTGGAGAGCTAGGAGAAACCGCGAAGTCAGTACACAGATCCCAAGAGAAAG 657
QY 703 GTGACCTGTGACATCCGCTCCGGGTTGAGCAGAGTACTGAGAGATGGGCGAGCTTG 762
DB 658 CAGACATGTGACATGACATGAGCTGGGGGTTGGGCTGAATCTGCACAGATGAGCTGCTG 717
QY 763 GAGCAGGCGTGGCATCCCGGCGCGCCCAAGGCGCTGGCGGAGCTGACGTGTTGGG 822
DB 718 CAGGGCAATGTCTTCTTAACAAGCAGGAGCCCACTTGAAGCGGCAAGTTTGAAGCCTTTAAC 777
QY 823 CAGGGCAACGAGTGTGCTGCTCAAGGAGACTGGGCTGTGTGTTGATCAATCAAGTTT 882
DB 778 CAGGGCAACCACTCTCAAGTCCCGGAGCTGGGCTCATATCTGTACATCAAGTTT 837
QY 883 TCAGAGCTCTCTATCTGAGCGCTTCTGGGCGACTGAGTCAAGTGGCGCCCTGTGAG 942
DB 838 TCTGAGCTACCTCACTGATGATTTCTGGCGTGAATCAATCAAGGCTCTTATTAGAG 897
QY 943 GCCCTGCGGGGCTGTCTTCTGACTGAGGCCCTTGCAGAGAGGCTGTGGCGGAGGCTGTT 1002
DB 898 GCATTTAAAGGTGTCTTCAATCAAGACTTCCCTCAAGCAAGCTGTGGGCCATGAGGCATC 957
QY 1003 CGCCTGTGCTCAGTGTGATGAGGCTGACTATGAGGCTGGCGGCGCCCTGTTGCTG 1062
DB 958 AAGCTGTGTAATGTAGAGAGGAGACTATGAGTGGGCCGAGCAAAATCTCTGAGG 1017
QY 1063 A 1063
DB 1018 A 1018
```

RESULT 8
 US-09-733-167A-4
 ; Sequence 4, Application US/09733167A
 ; Patent No. 6696547
 ; GENERAL INFORMATION:
 ; APPLICANT: Krammer, Peter
 ; APPLICANT: Krammer, Peter
 ; TITLE OF INVENTION: Protein for Regulation of Apoptosis
 ; FILE REFERENCE: 4121-120
 ; CURRENT APPLICATION NUMBER: US/09/733,167A
 ; CURRENT FILING DATE: 2000-12-08
 ; PRIOR APPLICATION NUMBER: PCT/DE99/01712
 ; PRIOR FILING DATE: 1999-06-08
 ; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
 ; PRIOR FILING DATE: 1998-06-08
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1142
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-09-733-167A-4

Query Match 10.7%; Score 201.2; DB 4; Length 1142;
 Best Local Similarity 52.3%; Pred. No. 5,6e-40;
 Matches 496; Conservative 0; Mismatches 443; Indels 9; Gaps 2;

116 CCCCAGGAAATGGGCTTATCCGGGTGACCCCGGCGGCTGTGGAGAGAGATAGTCC 175
 20 CCTTACATAGGCGGCTTAAGAGGCGGACCAAGTGTGCTCCGAAGAGCGTGGG 79
 176 TGAGCTACTAGGGATGCTGTGCTGACCGTATGTTGAGGTGTGGGCGGCACTGA 235
 80 AGCAGAGACTGGGCTTACAGCTTCAACCGCATGTTGACATGTGGACCCCACTAA 139
 236 CCGAGTGCAGCTGAGCTCTGCTGCTTGTGTGATGAGGCTCTGCGCGCCGCGAG 295
 140 CACAGAGATGTCGAGTGTCTTCTTCTTTGTGATGTTATGATGACCATGAC 199
 296 GCTTAGCCCGGCGGCGGCTTACAGCTCTGCTGAGCTGAGCGCGCGGCGAGT 355
 200 G---TGACTCATCCGAATGAGAGTGAATCTTATTGAGCACTGAGACGCGCAGGCGCT 256
 356 GCGGCGAGCAACTGCGGCTGCTGGGCAACTCTGGCGGTGCTGGCGCGCAGCACC 415
 257 GTGACGAGATTAATTTGCGCAGGTGCTGAGCTGCGCATATCACTGCGCATACT 316
 416 TGCTGCCGCACTGCGCGGAGCGGCGCGGCGGCTCTCCAGAAAGCTATAGTATG 475
 317 TGCTGCCCTAGTACTCTCAAGAAAGAGAGAGTGTGCTGCTGATCTGTAGACAAGT 376
 476 GCACCTTCAGCTTTCAAAGAGAGAGAGTGTGCTGCTGCTGCGGCGGCAAGCA 535
 377 ATCTGAGAAACATCAATTCGCTATGTGACCCCAAGAGCTGAGACCCAGAACGA 436
 536 GTTCTGCAATTTCTCAGCAGGCTGAGTGGAGAGAGGCTCCCCCAACCAAGGCGAGC 595
 437 GGCT-----TCCCAAGCCTCTTAAACAGTGTCTCCCAATATCTGTGTGTGCTGCC 490
 596 GCGGAGTGGGCGGCGCGGCGGAGTGTGTGTCAGACGCGCGGCGGAGAGGCGCCAGCG 655
 491 CCACTTGGGGTTCCTCAATGTGTAGTAAGGCGGCGGCGGAGGAGAACCACTTGGGA 550
 656 CACCCAGCAGCACTCAGAGCGCGGCGGCAAGCTTCTCTGAAGCAAAAGTACTGTGACA 715
 551 GCGCAGCAAAAACGCGGAGTGTGTGACACCAAGCCGGAAGGAAACAGACATGTATATA 610
 716 TCGGCTCCGGGTTTGAAGAGTACTGCGAGCATGGGCGGCACTTGAAGAGGCGGCTGG 775
 611 TCAAGCTCCAGATTCGAGCGGAAATATGCGCAGCATGAGACGCTCTGTGAAAGCAATGCT 670
 776 CATCCCGGCGGCGCGGCGGCGGCGGCGGAGCTGTGAGCTGTGTTGGGCAAGCCAGCCAG 835

Db 671 TCTCATAATAGCAGAGACCCCACTTGAGCGCGAGTTTGAGCCCTTAAACAGGCCAACACTA 730
 Qy 836 TGCTGCGCTCAAGGAGACCTGGGCTGTGTGTTGTGACATCAAGTCTAGAGCTGCT 895
 Db 731 TCTCAAGTCCCGGAGCTGGGCTCCATCATCTGTGACATCAAGTCTGTAGCTACCT 790
 Qy 896 ATCTGACGCTTCTGGGCGGAGCTTACCTGAGTGGCGGCTGCTGACAGGCCCTGGGCGG 955
 Db 791 ACCTGACGATCTTGGCGGAGCTTACCTTATATGCTCATTTATGAGGCACTGAAAGTG 850
 Qy 956 TGTCTGACTGAGGCGCTTGGAGGCTGTGTGCGGCGGAGGCTGTGCTGTGCTGA 1015
 Db 851 TCTTATCAGACAGACTCTTCAAGCAAGCTGTGGGCGCATGAAAGCATCAAGCTGTGTGA 910
 Qy 1016 GTGTGATGAGGCTGATGAGGCTGCGGCGGCGGCGGCTGTGCTGA 1063
 Db 911 ACGTGATGAGGAGTATGAGCTGTGGCGGCGGAGAAATCTCTGAGGA 958

RESULT 9
 US-08-859-167-5
 ; Sequence 5, Application US/08859167
 ; Patent No. 6037461
 ; GENERAL INFORMATION:
 ; APPLICANT: Alnemri, Emad S.
 ; APPLICANT: Fernandez-Alnemri, Teresa
 ; TITLE OF INVENTION: PADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
 ; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSES: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461-15
 ; STREET: One Liberty Place, 46th Floor
 ; CITY: Philadelphia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19103
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: WINDOWS
 ; SOFTWARE: Wordperfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/859,167
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Deluca, Mark
 ; REGISTRATION NUMBER: 33,229
 ; REFERENCE/DOCKET NUMBER: TJU-
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (215) 568-3100
 ; TELEFAX: (215) 568-3439
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1200 base pairs
 ; STRANDEDNESS: double
 ; TYPE: nucleic acid
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 32..988
 ; US-08-859-167-5

Query Match 10.7%; Score 201.2; DB 3; Length 1200;
 Best Local Similarity 52.3%; Pred. No. 5.7e-40;
 Matches 496; Conservative 0; Mismatches 443; Indels 9; Gaps 2;
 116 CCCCAGGAAATGGGCTTATCCGGGTGACCCCGGCGGCTGTGGAGAGAGATAGTCC 175
 24 CCTTACATAGGCGGCTTAAAGAGGCGGCAAGCCAGAGTGTGGCGCGGAGACGCTGGG 83

```

QY 176 TGACCTAAGGAGATGCTGTGCTTACCGTATGTTGAGTGTGGGCGGCAACTGA 235
DB 84 AGCAGAAACATGGGCTCTACAGCCTCCACCGCATGTTCCAGCATGTTGGGACCCACTAA 143
QY 236 CCGAGTGGCAGCTGAGAGCTCTGCGCTTTCTGTGATGAGGCTCCTGGGCGGCGGAG 295
DB 144 CACACAGAGATGTCGAGATGCTTCTCTCTTTTGTATGTTATGATGACATGAAC 203
QY 296 GCTTACCGCGGCGCGCGAGCGCTGAGCTCTGTGAGTGTGAGCGCGGCGGAGT 355
DB 204 G---TGACCTCATCCGAAATGAGACGTGACTTCTTATTTGAGACTGTGAGCGCGGCGGCT 260
QY 356 GCGGCGAGAGCAACTGCGGCTCTGCGGCGCACTCTGCGGCTGTGGCGCGGCGGAGCC 415
DB 261 GTGACGAGAGTACTTTCGCGCAGGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 320
QY 416 TGCTGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 475
DB 321 TGCTGCGGCTGCTTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380
QY 476 GCACCTTCAAGCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 535
DB 381 ATCTGAGAGAGAAATCAATTCGCTATGTGACCCCGAGAGCGCTCAGTGAACCGAAGCA 440
QY 536 GTTCTGCAATTTCTCAGCAGGCTGAGTGGGAGACAGGCTGCTCCCGCAACCAAGCGGCG 595
DB 441 GGCC-----TCCCGAGCCTCTTAAACAGTGTGCTCCCGCACTATCTGTGTGTGCTGCC 494
QY 596 GGGGAGAGTGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 655
DB 495 CCACTTGGGCTTCTCAATGTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 554
QY 554 CACCCAGCAGCAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 715
DB 555 GCGCAGGAGAAACCGCGGAGTCCGTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 614
QY 716 TCCGCTCCGCGGTTTCAAGAGAGTCTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAG 775
DB 615 TCAAGGCTCCGAGTTCGCGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 674
QY 776 CATCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 835
DB 675 TCTTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
QY 836 TGCTGCGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 895
DB 735 TCTTCAAGTCCGCGGAGCTGGGCTCCATCATCTGTGACATCAAGTTCTGTAGCTACCT 794
QY 896 ATCTGAGAGCTTCTGAGGCGGAGCTTACCTGAGTGGCGGCTGTGAGAGCGGCGGCGG 955
DB 795 ACCTGAGAGCTTCTGAGGCGGAGCTTACCTGAGTGGCGGCTGTGAGAGCGGCGGCGG 854
QY 956 TGTTCCTGAGTGGAGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1015
DB 855 TCTTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 914
QY 1016 GTCTGATGAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1063
DB 915 ACCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962

```

RESULT 10
US-09-109-273-5
; Sequence 5, Application US/09109273
; Patent No. 6063760

; GENERAL INFORMATION:
; APPLICANT: Alnemix, Emad S.
; APPLICANT: Fernandez-Alnemix, Teresa
; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
; OF MAKING THE SAME
; NUMBER OF SEQUENCES: 17

```

; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760r1s
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WINDOWS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/109,273
; FILING DATE:
; CLASSIFICATION:
; PRIORITY DATA:
; APPLICATION NUMBER: 08/859,167
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..988
; US-09-109-273-5
;
; Query Match 10.7%; Score 201.2; DB 3; Length 1200;
; Best Local Similarity 52.3%; Pred. No. 5, 7e-40;
; Matches 496; Conservative 0; Mismatches 443; Indels 9; Gaps 2;

```

QY	596	GGCGGAGTCTGGGGGCGGGCCAGTGGTGGGCGAAGCGGGGGAGAGGGGCCCGACCG	655
Db	495	CCACTTCGGGTTCCTCAAAATGTGTAAAGCGGCCGAGCCCGAGGGGAACACACTTTGGGA	554
QY	656	CACCCAGACAGCACTCAGAGCCCGCCAGACTTCTCTTGAAGGCAAAAGTACCTTGACA	715
Db	555	GCCAGCGAAAAACGCCGAAGTCGGTGAGAACCCAGACCCGAAGAAAAAGCAGACATGTGATA	614
QY	716	TCCGGCTCCGGGTTTCGAGCAGATCTCTCGAGCATGGCCAGCCCTTGGACAGGGCCGTGG	775
Db	615	TGAGGCTCCGAGTTTGGGGGGAAATCTGCGACGATGAAGACGGCTCTGGCAAGGCAATGCT	674
QY	776	CATCCCGGCGGCCCGACGGCGCTGGCGCGGACAGCTGATTTGGGAGAGCCACCGGAG	835
Db	675	TCTCAATAAAGAGGACCACCTTGAGCGCAGTTTGAAGGCTTTAAACAGGCGAACACTA	734
QY	836	TGCTGCGCTCAAGGAGCACTGGGCTCTGTGTGTTTGTGACATCAAGTTCTCAGAGCTCTCT	895
Db	735	TCTCAAGTCCCGGGACCTGGGCTCCATCTGTGAAATCAAGATTTCTGTAGCTCACCT	794
QY	896	ATTGGAAGCCTTCTTGGGGGCACTACTGAGTGGGGCCCTGTGAGGCCCTGGCGGGCG	955
Db	795	ACCTGAGGCATTTCTGGGGAATCAATTAATAGCTCAATTAATAGGCACTGAAAGGTG	854
QY	956	TGTTCTGACTGAGGCCCTGGGAGAGGCTGTGGGCCGGGAGGCTTTGCCCTGTGCTCA	1015
Db	855	TCTTCATACAGACTCTCTCAAGCAAGGCTGTGGCCATGAAGCCATCAAGGCTGCTGTGA	914
QY	1016	GTGTGAGTAGAGCTGACATTAAGGCTGGCCCGGCCCTGTGTGCTGA	1063
Db	915	ACGTGAGTAGAGACTATTAAGTCTGGGCCCAAGAACTCTCTTGAGA	962

RESULT 11
 US-09-276-993-5
 : Sequence 5, Application US/09276993
 : Patent No. 6207801
 : GENERAL INFORMATION:
 : APPLICANT: Alnemri, Emad S.
 : APPLICANT: Fernandez-Alnemri, Teresa
 : TITLE OF INVENTION: PADDD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
 : TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
 : OF MAKING THE SAME
 : NUMBER OF SEQUENCES: 17
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6207801r1s
 : STREET: One Liberty Place, 46th floor
 : CITY: Philadelphia
 : STATE: PA
 : COUNTRY: USA
 : ZIP: 19103
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: WINDOWS
 : SOFTWARE: Wordperfect
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/276,993
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/859,167
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Deluca, Mark
 : REGISTRATION NUMBER: 33,229
 : REFERENCE/DOCKET NUMBER: TJU-
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (215) 568-3100
 : TELEFAX: (215) 568-3439
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:

:	LENGTH:	1200	base pairs
:	TYPE:	nucleic acid	
:	STRANDEDNESS:	double	
:	TOPOLOGY:	both	
:	MOLECULE TYPE:	CDNA	
:	FEATURE:		
:	NAME/KEY:	CDS	
:	LOCATION:	32..988	
:	US-09-276-993-5		

Query Match	10.7%	Score 201.2;	DB 3;	Length 1200;
Best Local Similarity	52.3%;	Pred. No. 5,7e-40;		
Matches 496;	Conservative 0;	Mismatches 443;	Indels 9;	Gaps 2

OY	CCCCGGGAATGAGCGCTATCCGGATCGACCCCAGCCCCCGGTGCAGGAGAGATGATGCC	175
Db	24 CCTTACTGATGACGGGGCCTTAAGAAGCGGGCAAGCAGGTGTGGCCGAAGAGCTGGGG	83
OY	176 TGGACTACTACGGAGTGCTGTGCTTCACCGTATGTTGAGGTGTGGGCGGGCAATGA	235
Db	84 AGCAAGAACAATGGGCTTACAGCCTCCACCAGATTGTGACATGCTGGGCAACCACTAA	143
OY	226 CCGATGTGACCTGAGAGCTCCTGGCCCTTTCTGTGTGAATGAGAGCTTCCTGGCCCCGGAG	295
Db	144 CACACAGAGATGTCGAGTGTCTTCTCTTCTTTTGTGAATGTATATGACCATYGAC	203
OY	296 GCTTAGCCCCGGGCGCGAGCGGCTTAGAGCTCTGCTGTGAGACTGGAGCGCGCGGCACT	355
Db	204 G---TGACTCATTCGAAATGAGAGTGACTTTATTGGCACGTGAGGGCCAGGGCCCT	260
OY	356 GCGGAGAGCAACCTGCGGCTGCTGTGGGCAATCTCTGCGGTGCTGGCCCCCAAGACC	415
Db	261 GTGACGAGAGTAACTTTGCGCAGGTGTGCGAGCTGTGCGCATATCACTGCGCATGACT	320
OY	416 TGCTGCGCGCACTGGCGCGCAAGGGCGCGCGCCAGTGTCTCCAGAACGCTATAGCTATG	475
Db	321 TGCTGCTTACGTTAATCTTCAAAGAGACAGAGCTGTGTGCTGTATCTTTGTAGACAGT	380
OY	476 GCACCTTCAGCTCTTCAAAGAGACAGAGGATGCTGCGCGCTGCGGACGTCAAGCA	535
Db	381 ATCTGAGAGAAATCAATTCGTATGTGACCCCGAGAGCCCTCAGTACCCAGAACGA	440
OY	536 GTTCTGCAAAATTCACAGAGGTCAGTGGAGACAGGTCCTCCCCCAACCAAGCGGACG	595
Db	441 GGCC-----TCCCGAGCCCTCTTAAACAGTGCTCTCCCACTATCTGTGTGTGTGCC	494
OY	596 GCGGAGTCCGGGGCGGCGCCAGTGTGTGTCACAGCGCGCGCGAGAGAGGAGCCCGACG	655
Db	495 CCACCTTGGGGTCTTCAATGTGTGTAGAACGGCCAGCCCGAGGAGAACACACATTTGGGA	554
OY	656 CACCCAGACAGTCAAGTCAAGGCCCGCCAGACCTTCTCTGAAGGCAAAAGTACTGTGCA	715
Db	555 GCCAGCGCAAAAACCGCGGAATCGGTGACACAGACCCGAAAGAAAAAGCAGACATGTGATA	614
OY	716 TCGGGCTTCGGGTTTCGAGCAGAGTACTGCGAGCATGGGCGCAGCTTGGAGGAGGGCGTGG	775
Db	615 TCAAGGCTTCGAGTTTCGGCGGAATATGCGCAGCATGAGACGGCTCTGCAAGGCAATGCTT	674
OY	776 CATCCCGCGGCGCCCAAGCGCTGCGCGCGGCACTGTGACGTGTTTTGGGAGGCGCAACCGAG	835
Db	675 TCTCCAATTAAGCAGAGACCCACTTATGAGGCGCAGTTTGAAGCGCTTTAAACAGGCCAACTA	734
OY	836 TGTCTGCGCTCAAGGACCTTGGGCTCTGTGTGTTTGTGACATCAAGTTCTCAAGCTCTTCT	895
Db	735 TCTCCAAGTCCCGGGAACCTGGGCTTCATCACTGTGACATCAAGTTCTCTGAGCTCACTT	794
OY	896 ATCTGAGACGCTTTCGGGGCGCACTACCTGAGTGGCGCCCTGCTGACGCGCTGCGGGGCG	955
Db	795 AACTCGACGATTTCTGGCGAGACTCAKATTAATGTGCTATTATTAAGGCACTGAAAGGTG	854
OY	956 TGTTCCTGACTGAGGCGCTGTGAGAGAGGTGTGGCCGGGAGGCTGTTCGCTGTGTGTCA	1015
Db	855 TCTTCATCAACGATCTCTCAAGCAAGCTGTGGGCGATGAAGCCATCAAGCTGCTGTGTGA	914

```

QY      1016 GTGTGATGAGGCTGACTATGAGGCTGGCCGCGCCGCTGTGCTGA 1063
DB      915 AGGTGATGAGGAGGACTATGAGCTGGCGCGAAGAACTCTGAGGA 962

RESULT 12
US-09-723-450-5
; Sequence 5, Application US/09723450
; Patent No. 6576751
; GENERAL INFORMATION:
; APPLICANT: Alnemi, Emad S.
; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, Ar
; FILE REFERENCE: TJU2445
; CURRENT APPLICATION NUMBER: US/09/723,450
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/276,993
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/959,167
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: No. 6576751el Sequence
; LOCATION: (32)..(988)
; US-09-723-450-5

Query Match      10.7%; Score 201.2; DB 4; Length 1200;
Best Local Similarity 52.3%; Pred. No. 5,7e-40;
Matches 496; Conservative 0; Mismatches 443; Indels 9; Gaps 2;

QY      116 CCCCCGGAATGGCGCTATCCGGGTGACCCCGGCCCTGCTGGAGAGAGATGATGCC 175
DB      24 CCTCTGATGATGGCGGCTTAAAGAGCGGCGCAAGCGATGCGCCGAAGACGTGGGG 83

QY      176 TGACTACTACGGGATGCTGTCTCTTACCGTATGTTTCCAGAGGTGTGGGGGCACTGA 235
DB      84 AGGAGAACATGGGCTCTACAGCCTCCACCGCATGTTCGACATCGTGGGCAACCACTTA 143

QY      236 CCGAGTGGCAGCTGAGAGCTCTGCGCTTCTGTGTGATGAGAGTCTCCGTGGCGCCGCGAG 295
DB      144 CACACAGAGATGTCCAGAGTCTTCTCTTTTGTATGTATTTGATGACATGAAC 203

QY      296 GCTTAAACCCGCGCCGCGAGCGGCTTAAAGCTCTGCTGAGCTGAGCGCGCGGCGAGT 355
DB      204 G---TGAGTCAATCCGAAATGAGCGTGAATTCTTATTGGCACTGAGCGCGAGGCGCT 260

QY      356 GCGGCGAAGAACCTGCGGCTGTGGGGCACTCTCGCGCGTGTGGCCCGCAAGACC 415
DB      261 GTACACAGAGTAACTTTCGACAGTCTGTGCGACTGATCATCATCTGCGCATGACT 320

QY      416 TGTGCGCAGCTGGCGCGAAGCGGCGCGGCGAGTGTCTCAGAACGCTATAGCTATG 475
DB      321 TGTGCTCTTACGTTACTCTTAAAGAGAGAGAGCTGTGTGCTCTTGTGACAGT 380

QY      476 GCACCTTCAGCTTTCAAGAGAGACAGAGGAGTGTCCGCTGCGCGAGTCAAGCA 535
DB      381 ATCTGAGAGAAACATCAATTCGTATGTACCCCGAGAGCCCTCAGTGAACCGAACGA 440

QY      536 GTTCTGCAATTTCTCAGCAGGGTCAGTGGAGACAGGCTCCCCCAACCAAGCGGACG 595
DB      441 GGGC-----TCCCGAGCCCTTAAACAGTGTCTCCCACTATCTGTGTGTGCTCC 494

QY      596 GGGGAGTGGGGCGCGCCAGCGAGTGTGTGTCAGACGCGCGCGAGAGGGGGCCCGAGCG 655
DB      495 CCACTTGGGTTCTCAATATGTATGTAAGCGGCGCAGCCGAGGAGAACCACTTGGGA 554

```

```

QY      656 CACCCGAGCAGATCAGAGCCCGCAGACTTCTCTGAAAGCAAAAGTACCTGTGACA 715
DB      555 GCCAGGAAAACGCCGGAAGTGTGACACAGACCCGAGAAAGACAGACATGTGATA 614

QY      716 TCCGCTCCGGGTTGAGCAGAGTACTGCGACATGGGCGACGCTTGGAGCAGGGCGTGG 775
DB      615 TCAGGCTCCGAGTTGGGGGAGATCTGTGACATGAGCGGCTCTGCAAGCAATGTCT 674

QY      776 CATCCGCGCGCCCGCAGCGCTGCGCGCAGCTGAGACGTGTTTGGGCGAGCCGCGAG 835
DB      675 TCTCAATAAGAGAGAGCCACTTGAAGCGCCATTTTGAAGCGCTTTAAACAGGCAACACTA 734

QY      836 TGTGCGCTCAAGGAGACTGGGCTGTGTGTTTGTGACATCAAGTTCTGAGCTTCTCT 895
DB      735 TCTCAAGTCCCGGAGCTGGGCTCATCATCTGTGACATCAAGTCTCTGTGACTCACC 794

QY      896 ATCTGAGCGCTTTCGGGCGACTACCTGAGTGGGCGCCCTGTCGAGCGCCGCGGGCG 955
DB      795 ACCTGACGCAATTCGTGGCAGACTATTAATGTCTCATTTATTAAGGCACTGMAAGGTG 854

QY      956 TGTTCCTGACTGAGGCGCTGCGAGAGGCTGTGGCGGAGGCTGTTCGCTGTGCTGA 1015
DB      855 TCTTCATCAACAGACTCTCTCAAGCAAGCTGTGGCCATGAAGCCATCAAGCTGTGCTGA 914

QY      1016 GTGTGATGAGGCTGACTATGAGGCTGGCGCGCCGCTGTGCTGA 1063
DB      915 AGGTGATGAGGAGGACTATGAGCTGGCGCGAAGAACTCTGAGGA 962

RESULT 13
US-09-733-167A-8
; Sequence 8, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Truncated segment of the DNA sequence encoding human DEBD, compri
; OTHER INFORMATION: sing nucleotides 352-981 of seq id no 2.
; US-09-733-167A-8

Query Match      8.3%; Score 156.4; DB 4; Length 630;
Best Local Similarity 60.2%; Pred. No. 5,5e-29;
Matches 259; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY      634 CCGCGAGAGGGGCCCCCAGCCGACCCGACGACGACTAGAGCCCGCAGACCTTCTCT 693
DB      178 CGAGGAGAGCCACATTTGGAGCCGACGAAAACGCGGAAATGACTGACCCAGATCCC 237

QY      694 GAAGGAAAGTACCTGTGACATCCGCTCCGGGTTTCAGCAGAGTACTGCGAGATGGG 753
DB      238 MAGAGAAACAGATATATACATCAGACTGCGGGTTCCGGGCTGAATCTGCCAGATGAG 297

QY      754 CCAAGCTTGAAGCAGGCGCTGCAATCCCGCGCGCCCGCAGCGCTGCGCGGACGTGAG 813
DB      298 ACTGCTCTCAGGAGCAATGTCTTCTTAACAAGACAGAACCACTTGAAGCGCGAGTTTGA 357

QY      814 GTTTTGGGCAAGCCACCGCAAGTCTGCGCTCAAGGGAACCTGGGCTCTGTGTGTGAC 873

```


Accession	Sequence	Position
Db	CGCTTTAACGAGCCACACCATCTCTCAAGTCCCGGACCTGGCTCCATCATCTGTGAC	417
Qy	ATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTCTGGGCGCATCACTGAGTGGCGC	933
Db	ATCAAGTTCTCTGACCTACCTACTCGATGATTTCTGGCGGACTAATCATCATGCTCT	477
Qy	CTGCTGACAGCCCTTCGCGGCGCTGTCTTCCTGACCTGAGCCCTTCGAGAGGCTGTGGCGCG	993
Db	TTATTAGAGGCACTTAAAGTGTCTTCATCAACAGACTCCCTCAACCAAGCTGTGGGCCAT	537
Qy	GAGCGTGTTCGCTGCTGATCAGTGTGATGAGCGTGACTTAGAGGCTGGCGGCGCGC	1055
Db	GAAAGCATCAACGCTCTCTGTAAATTTAAGACGAGGACTAAGACTGGGCCGACAGAAA	597
Qy	CTGTGTGCTGA 1063	
Db	CTCTCTGAGGA 607	

```

RESULT 14
US-09-733-167A-7
Sequence 7, Application US/09733167A
Patent No. 6696547
GENERAL INFORMATION:
APPLICANT: Peter, Marcus
APPLICANT: Krammer, Peter
TITLE OF INVENTION: Protein for Regulation of Apoptosis
FILE REFERENCE: 4121-120
CURRENT APPLICATION NUMBER: US/09/733.167A
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: PCT/DE99/01712
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
PRIOR FILING DATE: 1998-06-08
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 342
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Truncated segment of the DNA sequence encoding human BADD, comprising
US-09-733-167A-7

```

Query Match	4.7%	Score 88.4	DB 4	Length 342
Best Local Similarity	57.3%	Pred. No. 3e-12		
Matches	180	Conservative	0	Mismatches 13
				Indels 3
				Gaps 1
QY	150	CCCGTGCTGGAGGAGATGATGATGCTTGGACTACACGGATGCTGTGCTTCAACCGTAT	209	
Db	27	CCAGGTGTGGCCAGAGAGCATGTGTGACAGAGAACACGGGCTGTACAGCTTCGACCCGAT	86	
QY	210	GTTGAGGTGTGTGGCGGGCACTGACCGAGTGCAGCTGGAGCTCCTGGCCTTTCTGCT	269	
Db	87	GTTTGACATCGTGGGCACTCATCTGACACACAGAGATGTGGCGGTCTTTTCTCTCTT	146	
QY	270	GGATGAGGCTCCTGGCGCGCGCGGAGGCTTATGCCCGGCGCGGACCGGCTTACAGAGCTCT	329	
Db	147	TGTTGATATCTATTGATGTACACAGACGTGACCTCAT--CCGAAATGAGCTGTACTTTCT	203	
QY	330	GCTGAGCTGAGCGCGCGCGGAGTGCAGGAGAGCAACCTGTGGGCTGTCTGGGGCACT	389	
Db	204	ATTGGCACTGGAGCGCGCGGAGGCGCGGTGTGATGAAGTAACTTTTCGCCAGGTGTGACGCT	263	
QY	390	CCTGGCGCTGCTGAGCGCGCGCAAGACTGTCTGTCCGCACTGTGCGCGGCAAGCGGCGCGGCTC	449	
Db	264	GCTGGGCACTCATCACTCCGCCACGACCTGTCCCTCACTCAACCTCAAGAGAGACGGGCG	323	
QY	450	AGTGTCTCCAGAAC	463	
Db	324	TGTGTGCCCTATTC	337	

RESULT 15
 US-08-232-463-14
 : Sequence 14: Application US/08232463
 : Patent No. 5670367
 : GENERAL INFORMATION:
 : APPLICANT: DORNER, F.
 : APPLICANT: SCHEIFLINGER, F.
 : APPLICANT: FALKNER, F. G.
 : TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 : NUMBER OF SEQUENCES: 52
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Foley & Lardner
 : STREET: 1800 Diagonal Road, Suite 500
 : CITY: Alexandria
 : STATE: VA
 : COUNTRY: USA
 : ZIP: 22313-0299
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/232,463
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/935,313
 : FILING DATE:
 : APPLICATION NUMBER: EP 91 114 300.6
 : FILING DATE: 26-AUG-1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BENT, Stephen A.
 : REGISTRATION NUMBER: 29,768
 : REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (703) 836-9300
 : TELEFAX: (703) 683-4109
 : TELEX: 899149
 : INFORMATION FOR SEQ ID NO: 14:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 7218 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : CLONE: pTZgpc-F18
 :
 : US-08-232-463-14

Query Match	3.0%;	Score 56.2;	DB 1;	Length 7218;
Best Local Similarity	4.2%;	Pred. 0.00094;		
Matches	16;	Conservative 218;	Mismatches 151;	Indels 0; Gaps 0;
Qy	1125	TCCCACTCCAGCTCCGGGCCACTCTCTCTGGAGGACGACCACTCTACCCCTTGA	1184	
Db	1077	YY	1136	
Qy	1185	CAGCCCTCCACAGATGTGGGCTCTGAGGCTTAACATTTCCAGCTAGTTCTTC	1244	
Db	1137	YY	1196	
Qy	1245	CCAGACTCTCTCTACCCCGCAGGTGTGCCCTTACCTCGGAGGCGGGGCTGGGCTG	1304	
Db	1197	YY	1256	
Qy	1305	TATCTCAGAGGAGGAGGACAGCTACACTCACCAAGGCCCCCTGCAATTGTATC	1364	
Db	1257	YY	1316	
Qy	1365	TCTGATCTTGGGCTGTCTGCACTGTCAAGGTGCACACTCGCTCATCTCACACTGCC	1424	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 23, 2005, 05:31:52 ; Search time 1090.55 Seconds
(without alignments)
10221.347 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 1883
Sequence: 1 aggcgtatcatcagagaag.....ctgaaaggscatagctg99 1883

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1883	100.0	1883	5	AAF27408
2	1812.4	96.3	1979	12	ADQ86891
3	1787.2	94.9	2045	3	AAAS5790
4	1787.2	94.9	2045	4	AAH3283
5	1775.2	94.3	2044	3	AAH18296
6	1681.4	89.3	1924	6	AAH40080
7	1681.4	89.3	1924	6	AAH59062
8	1630.4	86.6	1966	6	AAH52603
9	1339	71.1	1570	6	AAH99646
10	1180.8	62.7	1230	6	AAH94362
11	979.4	52.0	981	10	ADC79259
12	946.4	50.3	1106	6	ABL39693
13	909	48.3	909	5	AAF27407
14	902.4	47.9	1084	6	ABL39692
15	719	38.2	847	11	AD130856
16	705.4	37.5	857	6	ABZ70266
17	628	33.4	1067	4	AA158961
18	628	33.4	1067	5	ADQ99183
19	628	33.4	1067	9	ADB48943
20	616	32.7	1042	4	AA160747

21	615	32.7	626	4	AAH07925	AAH07925 Human cDN
22	457.2	24.3	603	5	AAF93937	AAf93937 Primer SP
23	365.6	19.4	11084	12	ADQ18808	Adq18808 Human SOF
24	355.8	18.9	381	4	AAH12893	AAH12893 Human cDN
25	346.6	18.4	360	2	AAH25404	AAH25404 Human gen
26	302.8	16.1	451	9	ACH41209	ACH41209 Human fce
27	302.8	16.1	484	9	ACH25056	ACH25056 Human adu
28	301.4	16.0	303	6	AAH40075	AAH40075 Human DED
29	301.4	16.0	303	6	AAH59057	AAH59057 Human DED
30	296.8	15.8	441	6	ABL80091	ABL80091 Human ova
31	206.2	11.0	1139	3	AAZ43923	AAZ43923 Human DED
32	206.2	11.0	1830	4	AAH15765	AAH15765 Human cDN
33	206.2	11.0	2261	4	AAH14367	AAH14367 Human cDN
34	206.2	11.0	2261	6	ABK46133	ABK46133 cDNA enco
35	204.6	10.9	1045	2	AAH41337	AAH41337 Human FLA
36	204.6	10.9	1045	5	AAH43203	AAH43203 Human FLA
37	204.6	10.9	2079	5	ABV24761	ABV24761 Human pro
38	201.2	10.7	1142	3	AAZ43924	AAZ43924 Murine DE
39	201.2	10.7	1200	2	AAV74138	AAV74138 Mouse FLA
40	201.2	10.7	1200	6	AAH43204	AAH43204 Mouse FLA
41	168	8.9	594	13	ADQ55421	ADQ55421 Mouse mit
42	129.4	6.9	769	10	ADD34431	ADD34431 Mouse mit
43	124.8	6.6	525	10	ADD34432	ADD34432 Mouse mit
44	122.2	6.5	553	9	ACH23692	ACH23692 Human adu
45	121.2	6.4	422	8	ABZ18661	ABZ18661 Group III

ALIGNMENTS

RESULT 1	AAH07925
ID	AAH07925
XX	AAH07925 standard; cDNA; 1883 BP.
XX	AAH07925
XX	24-APR-2001 (first entry)
XX	Human apoptosis-associated factor NT2RM1000558 cDNA, SEQ ID NO:3.
XX	Human apoptosis-associated factor; NT2RM1000558; death effector domain;
KW	DED; caspase family cleavage domain; pro-apoptotic; drug screening;
KM	cell proliferation; ischemic disease; chronic viral disease; ss.
XX	
OS	Homo sapiens.
XX	
XX	WO200104300-A1.
XX	18-JAN-2001.
XX	
XX	06-JUL-2000; 2000WO-JP004516.
XX	
XX	08-JUL-1999; 99JP-00194179.
XX	18-OCT-1999; 99US-0159586P.
XX	
PA	(HELI-) HELIX RES INSTR.
XX	
PI	Ota T, Isegai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
XX	WPI; 2001-138348/14.
DR	P-PDB; AAB60387.
XX	
PT	Polynucleotide encoding an apoptosis-associated factor protein with death
PT	effector domain and caspase family-cleavage domain, useful in regulating
XX	diseases with cell proliferation.
XX	
XX	Claim 3; Page 44-47; 53p; Japanese.
CC	The invention relates to a novel human apoptosis-associated factor
CC	(AAB60386, AAB60387), designated NT2RM1000558, which contains a death
CC	effector domain (DED), and a caspase family cleavage domain and is capable
CC	of inducing apoptosis in cells. The invention also relates to nucleic
CC	acids encoding the protein (AAH07925, AAF27407, AAF27408); variants of the protein

CC (particularly dominant negative variants); vectors and host cells
CC comprising a nucleic acid which encodes an apoptosis- associated factor
CC of the invention; the recombinant production of the protein; an antibody
CC against the protein; and methods of screening for compounds which can
CC regulate apoptosis. The apoptosis-related factor is useful in regulating
CC diseases associated with cell proliferation and in screening drug
CC candidates e.g., for regulating cell proliferation or cell death in
CC ischaemic diseases and chronic viral diseases. The present sequence
CC represents a full-length cDNA encoding the human apoptosis-associated
CC factor NT2RM100558

XX Sequence 1883 BP; 343 A; 595 C; 588 G; 357 T; 0 U; 0 Other;

Query Match 100.0%; Score 1883; DB 5; Length 1883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCGTAATATATGAGAAAGTGCAGAAAGATCCAAACAGAGGCTGCGCCCTGCCCG 60
DB 1 AGGCGTAATATATGAGAAAGTGCAGAAAGATCCAAACAGAGGCTGCGCCCTGCCCG 60
QY AGGAGTCATCGAGCGCCAGATCTGAGCTTGTTCGCGCTCCCTCCCGC 120
DB AGGAGTCATCGAGCGCCAGATCTGAGCTTGTTCGCGCTCCCTCCCGC 120
QY 121 GGAATGCGCTATCCGGGTGACCCCGGCCCGCTGCGAGAGAGATGATGCTTGAC 180
DB 121 GGAATGCGCTATCCGGGTGACCCCGGCCCGCTGCGAGAGATGATGCTTGAC 180
QY 181 TACTACGGAGATGCTGCTGCTTCAACCTATGTTTGAAGTGTGGGCGGCAATGACCGAG 240
DB 181 TACTACGGAGATGCTGCTGCTTCAACCTATGTTTGAAGTGTGGGCGGCAATGACCGAG 240
QY 241 TGGAGCTGAGAGCTCTGCGCTTCTGCTGATGAGGCTCTGCGCCCGCGAGGCTTA 300
DB 241 TGGAGCTGAGAGCTCTGCGCTTCTGCTGATGAGGCTCTGCGCCCGCGAGGCTTA 300
QY 301 GCCCGGCGCCGAGCGGCTGAGGCTCTGCTGAGCTGAGCGCGCGGCGAGTGGCGC 360
DB 301 GCCCGGCGCCGAGCGGCTGAGGCTCTGCTGAGCTGAGCGCGCGGCGAGTGGCGC 360
QY 361 GAGAGAACTGCGGGGCTGCTGGGGGCAATCTCGCGGCTGCTGGCCCGGCAACGCTGCTG 420
DB 361 GAGAGAACTGCGGGGCTGCTGGGGGCAATCTCGCGGCTGCTGGCCCGGCAACGCTGCTG 420
QY 421 CGCAGCTGCGCGGCAAGCGGCGCGGCAAGTCTTCCAGAACGCTATATGAGCAAC 480
DB 421 CGCAGCTGCGCGGCAAGCGGCGCGGCAAGTCTTCCAGAACGCTATATGAGCAAC 480
QY 481 TCCAGCTTTCAAAAGAGACAGAGGTAAGTGCCTGCGCGCTGCGGAGTCAAGCAATTTCT 540
DB 481 TCCAGCTTTCAAAAGAGAGAGAGGTAAGTGCCTGCGCGCTGCGGAGTCAAGCAATTTCT 540
QY 541 GCAAAATTTTCAGAGAGGTCAGTGGAGAGAGGCTCCCCCGCAACCAAGGCGAGCGGCGG 600
DB 541 GCAAAATTTTCAGAGAGGTCAGTGGAGAGAGGCTCCCCCGCAACCAAGGCGAGCGGCGG 600
QY 601 AGTCGCGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 AGTCGCGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CAGCAGCACTGAGAGCGCGGCAAGCTTCTTGAAGGCAAGTGAACCTGATCCCG 720
DB 661 CAGCAGCACTGAGAGCGCGGCAAGCTTCTTGAAGGCAAGTGAACCTGATCCCG 720
QY 721 CTTCGCGGCTTGAAGAGATCTGCGAGCATGAGCGCTTGAAGAGCGGCGGCTGATCC 780
DB 721 CTTCGCGGCTTGAAGAGATCTGCGAGCATGAGCGGCTTGAAGAGCGGCGGCTGATCC 780
QY 781 CGGCGGCGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 840
DB 781 CGGCGGCGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 840

QY 841 CCTCAAGGAGCCTGCGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTATCTG 900
DB 841 CCTCAAGGAGCCTGCGCTCTGTGTTGTGATCATCAAGTTCTCAGAGCTTCTATCTG 900
QY 901 GAGCGCTTCTGCGGAGACTACCTGAGTGGCGGCTGCGAGAGGCTGCGGCGGCTGCTG 960
DB 901 GAGCGCTTCTGCGGAGACTACCTGAGTGGCGGCTGCGAGAGGCTGCGGCGGCTGCTG 960
QY 961 CTGACTGAGGCGGCTGCGAGAGGCTGCGGCGGCTGCGGAGGCTGCTGCGGCTGCTG 1020
DB 961 CTGACTGAGGCGGCTGCGAGAGGCTGCGGCGGCTGCGGAGGCTGCTGCGGCTGCTG 1020
QY 1021 GATGAGGCTGATATGAGGCTGCGGCGGCTGCGGAGGCTGCTGATGAGAGAGGAGG 1080
DB 1021 GATGAGGCTGATATGAGGCTGCGGCGGCTGCGGAGGCTGCTGATGAGAGAGGAGG 1080
QY 1081 CGGCGGCGGAGAGGCGGCTGCTGATCCAGAGCTGCGAGAGTGTATCCACTTCAAGTCT 1140
DB 1081 CGGCGGCGGAGAGGCGGCTGCTGATCCAGAGCTGCGAGAGTGTATCCACTTCAAGTCT 1140
QY 1141 CCGGGGCACTTCTCTGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
DB 1141 CCGGGGCACTTCTCTGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
QY 1201 ATGTGGGCTCTGAGGCTTAAACATTTTCAAGCTGAGTTCTTCCAGACTCTCTAC 1260
DB 1201 ATGTGGGCTCTGAGGCTTAAACATTTTCAAGCTGAGTTCTTCCAGACTCTCTAC 1260
QY 1261 CCAGAGTGGCGGCGGCTGAGGCTGCGGAGGCGGAGGCGGAGGCTGATCTCAGAGAGG 1320
DB 1261 CCAGAGTGGCGGCGGCTGAGGCTGCGGAGGCGGAGGCGGAGGCTGATCTCAGAGAGG 1320
QY 1321 GGCAGAGCTACACACTCAACAAAGGCGGCGGCTGCAATTTATCTCTGATCTTGGGCTG 1380
DB 1321 GGCAGAGCTACACACTCAACAAAGGCGGCGGCTGCAATTTATCTCTGATCTTGGGCTG 1380
QY 1381 CTGCACTGTCAAGGTGACACACTGCTGATGCTCACTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1381 CTGCACTGTCAAGGTGACACACTGCTGATGCTCACTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 CTGGGCTGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 CTGGGCTGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 TCAGAGCTCTAATTTGACCAACCAACCTGAGGCTTCAAGCAATCACTGAGGCACTGGA 1560
DB 1501 TCAGAGCTCTAATTTGACCAACCAACCTGAGGCTTCAAGCAATCACTGAGGCACTGGA 1560
QY 1561 GCTGGGCTGACATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1561 GCTGGGCTGACATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 1621 GGCAGGCTTCAATTTACAGACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
DB 1621 GGCAGGCTTCAATTTACAGACCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1681 ACATGACTTGAACCTGATGATGACCTGAGAGTGTACATGAGGAGTGTATGAGAGATCA 1740
DB 1681 ACATGACTTGAACCTGATGATGACCTGAGAGTGTACATGAGGAGTGTATGAGAGATCA 1740
QY 1741 TACCAAGGCTTACTGTTGACATGAGGCGGCAACCAAGTGAACGCACTTCTTGGAAAG 1800
DB 1741 TACCAAGGCTTACTGTTGACATGAGGCGGCAACCAAGTGAACGCACTTCTTGGAAAG 1800
QY 1801 GGAATGCAAAAGGCTTGGGCGGATGAGAAAGACCTTTTACAATGATACCAATTTAACT 1860
DB 1801 GGAATGCAAAAGGCTTGGGCGGATGAGAAAGACCTTTTACAATGATACCAATTTAACT 1860
QY 1861 GCCCTGGAAGGAGCATAGGCTGAG 1883
DB 1861 GCCCTGGAAGGAGCATAGGCTGAG 1883

RESULT 2
 ADQ86891
 ID ADQ86891 standard; cDNA; 1979 BP.
 XX
 AC ADQ86891;
 XX
 DT 07-OCT-2004 (first entry)
 XX
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #3766.
 XX
 DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KM cancer; cell proliferative disorder; gene; ss.
 XX
 OS Homo sapiens.
 OS
 XX
 PN MO2004060270-A2.
 XX
 PD 22-JUL-2004.
 XX
 PF 15-OCT-2003; 2003MO-US029126.
 XX
 PR 18-OCT-2002; 2002US-0418988P.
 XX
 PA (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 PI Wu TD, Zhou Y;
 XX
 DR WPI, 2004-534300/51.
 XX
 PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 PS Claim 1; SEQ ID NO 3766; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identity to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising the presence of a protein in a sample suspected of containing the protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a

CC human TAT cDNA sequence from the present invention.
 XX
 SQ Sequence 1979 BP; 365 A; 617 C; 627 G; 370 T; 0 U; 0 Other;
 Query Match 96.3%; Score 1812.4; DB 12; Length 1979;
 Best Local Similarity 96.9%; Pred. No. 0;
 Matches 1882; Conservative 0; Mismatches 1; Indels 59; Gaps 1;

1 AGCGTAAATATGAGAGAGTGGCCAGAAAGATCCAAACAAAGAGTGGCGGCGTGGCCG
 38 AGCGTAAATATGAGAGAGTGGCCAGAAAGATCCAAACAAAGAGTGGCGGCGTGGCCG
 61 AGGAGTCAATGAGAGCGCAGAAATCTGGCCGGGTTCTGAGCTTGTCCGCTCCCTCCCG
 98 AGGAGTCAATGAGAGCGCAGAAATCTGGCCGGGTTCTGAGCTTGTCCGCTCCCTCCCG
 121 GGAATGGCGCTATCCGGGTTGAGACCCCGGCGGCTGGGAGAGATGATGTCCTGGAC
 158 GGAATGGCGCTATCCGGGTTGAGACCCCGGCGGCTGGGAGAGATGATGTCCTGGAC
 181 TACTACGGAGTGGCTGCTGCTTCAACCGTATGTTGAGAGTGGGCGGCGCACTGACGAG
 218 TACTACGGAGTGGCTGCTGCTTCAACCGTATGTTGAGAGTGGGCGGCGCACTGACGAG
 241 TGGAGCTGAGAGCTCTGAGCTTTCTGATGAGAGCTCTGAGCGCCGCGGAGGCTTA
 278 TGGAGCTGAGAGCTCTGAGCTTTCTGATGAGAGCTCTGAGCGCCGCGGAGGCTTA
 301 GCCCGGCGCGCGCGCGCTGAGAGCTCTGAGAGCTCTGAGCGCCGCGGAGGCTTA
 338 GCCCGGCGCGCGCGCGCTGAGAGCTCTGAGAGCTCTGAGCGCCGCGGAGGCTTA
 361 GAGAGCAACCTGCGGCTGCGGCGGCGCGAGCTCTGAGAGCTCTGAGCGCCGCGGAGGCTTA
 398 GAGAGCAACCTGCGGCTGCGGCGGCGAGCTCTGAGAGCTCTGAGCGCCGCGGAGGCTTA
 421 CCGACCTGCGGCGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 458 CCGACCTGCGGCGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 481 TCGAGCTCTTCAAAAG
 518 TCGAGCTCTTCAAAAG
 541 GCAATTTTCAG
 578 GCAATTTTCAG
 601 AGTGGGCGCGCGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 638 AGTGGGCGCGCGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 661 CAGCAGAGTCAAGAGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 698 CAGCAGAGTCAAGAGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 721 CTCGGGTTTCAG
 758 CTCGGGTTTCAG
 781 CCGCGCGCGCGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 818 CCGCGCGCGCGCGCGCGCGCGCGAGTCTTCAGAAAGCTATAGCTATGAGCACC
 841 CCGTCAAG
 878 CCGTCAAG
 901 GAGCGCTTCGGGCGCGAGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 938 GAGCGCTTCGGGCGCGAGTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 961 CTGACTGAG

Db	998	CTGACTAGAGGCCCTGCGAGAGGCTGTGTGGCCGGGAGGCTGTGTCCGCTGCTGTGTCAGTGTG	1057
Qy	1021	GATGAGCTGACTATGAGGCTGGCCGCGCGCCCTGTTGCTGATGAGAGAGAAAGGGGG	1080
Db	1058	GATGAGGCTGACTATGAGGCTGGCCGCGCGCGCCCTGTTGCTGATGAGAGAGAAAGGGGGG	1117
Qy	1081	CGGGGCCCCGAGAGAGGCGCTCCTGATTCGAGGACTGCGAGAGATTGATCCCACTCCAACTG	1140
Db	1118	CGGGGCCCCGAGAGAGGCGCTCCTGATTCGAGGACTGCGAGAGATTGATCCCACTCCAACTG	1177
Qy	1141	CGGGGCACTTCTCTCTGGAGAGACGACATCTCTA-----	1176
Db	1178	CGGGGCACTTCTCTCTGGAGAGAGACATCTCTA-----	1237
Qy	1177	-----CCCTTGACAGCCCTCCCAAGAA	1201
Db	1238	ATCTTTGAGACTGCGACAGGACCGGGGACAGAGGCCCTTGACAGGCCCTCCCAAGAA	1297
Qy	1202	TGTGGGCTCTGAGGCGCTTAAACATTTCCAGCTGAGTTTCTTTCCTCCAGACTCTCTTACC	1261
Db	1298	TGTGGGCTCTGAGGCGCTTAAACATTTCCAGCTGAGTTTCTTTCCTCCAGACTCTCTTACC	1357
Qy	1262	CCAGGTGTCGCCCTTAGCTCTCCGAGGCGGGGGGTGGGCGCTGATATCTCAGAAAGGAGG	1321
Db	1358	CCAGGTGTCGCCCTTAGCTCTCCGAGGCGGGGGGTGGGCGCTGATATCTCAGAAAGGAGG	1417
Qy	1322	GCACAGCTACACACTCACAAAGGCCCCCTGCACATTTGATCTCTGATCTTGGGCTGTC	1381
Db	1418	GCACAGCTACACACTCACAAAGGCCCCCTGCACATTTGATCTCTGATCTTGGGCTGTC	1477
Qy	1382	TGCATCTGCACAGGTGCACACACTGCTCATGCTCACACTGCGCTGTGAGATCTTCCC	1441
Db	1478	TGCATCTGCACAGGTGCACACACTGCTCATGCTCACACTGCGCTGTGAGATCTTCCC	1537
Qy	1442	TGGGCGCTGTGCGCTGGCGCTGCTCCAGACACACTTCTTGGGCTTAAAGGACTCTGCT	1501
Db	1538	TGGGCGCTGTGCGCTGGCGCTGCTCCAGACACACTTCTTGGGCTTAAAGGACTCTGCT	1597
Qy	1502	CAGGACCTCTAATTGACACAGAACCAACCTGGGCTTCAGCCACATCAGTGGGCACTGGAG	1561
Db	1598	CAGGACCTCTAATTGACACAGAACCAACCTGGGCTTCAGCCACATCAGTGGGCACTGGAG	1657
Qy	1562	CTGGGGTGCACATGGGGGCTGTGCTACCTTGCCACACATCTCCAGCCAGCGAGGCGCTG	1621
Db	1658	CTGGGGTGCACATGGGGGCTGTGCTACCTTGCCACACATCTCCAGCCAGCGAGGCGCTG	1717
Qy	1622	CCGAGCTTCATTTTACAGACCTGACTCTCCCTCACTTCCCGGCTGGTCCAGAGCTGAA	1681
Db	1718	CCGAGCTTCATTTTACAGACCTGACTCTCCCTCACTTCCCGGCTGGTCCAGAGCTGAA	1777
Qy	1682	CATAGACTTGCATTTGATGTCACTGTGAGTGCATATGGAGTGTATGCGACATCAT	1741
Db	1778	CATAGACTTGCATTTGATGTCACTGTGAGTGTCACTGTGAGTGTATGCGACATCAT	1837
Qy	1742	ACCAAGGCTTACTGTGGACATGTGGGCGCAAAACAGTAAACAGCCACTTCTTGAAAG	1801
Db	1838	ACCAAGGCTTACTGTGGACATGTGGGCGCAAAACAGTAAACAGCCACTTCTTGAAAG	1897
Qy	1802	GAATTCGAAGGCTTTGGGGGTGATGGAAGAAAGCTTTTACAAATGATACCAATTAACTG	1861
Db	1898	GAATTCGAAGGCTTTGGGGGTGATGGAAGAAAGCTTTTACAAATGATACCAATTAACTG	1957
Qy	1862	CCCTGGAAGGCACTAGGTGGG 1883	
Db	1958	CCCTGGAAGGCACTAGGTGGG 1979	

XX	28-FEB-2001	(first entry)
DT	Apoptosis related gene 1 clone HLDK36.	
XX	Cycostatic; antiinflammatory; immunosuppressive; antisclerotic; cardiant;	
XX	vinculid; anti-AIDS; vasotropic; anti-ischaemic; antiparkinsonian; ss;	
KW	anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;	
KW	colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;	
KW	autoimmune disorder; multiple sclerosis; viral infection.	
XX	Homo sapiens.	
XX	WO200056752-A2.	
PN	28-SEP-2000.	
PD	15-MAR-2000; 2000WO-US006642.	
PF	24-MAR-1999; 99US-0126018P.	
PR	17-JUN-1999; 99US-0139638P.	
PR	18-AUG-1999; 99US-0149449P.	
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Ruben SM, Ni J, Young PA.	
PI	WPI; 2000-587660/55.	
DR	P-PSDB; AAB15551.	
DR	Nucleic acid encoding human apoptosis associated protein, useful for the	
PT	prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's	
PT	disease, inflammation and ischemic injury.	
XX	Claim 1; Page 244-245; 273pp; English.	
PS	The invention relates to the isolation of genes encoding a human	
XX	apoptosis-related proteins. The nucleotide sequences AAB15570-A95798	
CC	encode the human apoptosis related proteins AAB15551-B15559. The genes	
CC	can be used to generate fusion proteins by linking to the gene for the	
CC	human immunoglobulin G Fc (IgG Fc) portion (AAB15799) for increasing the	
CC	stability of the fusion protein as compared to the human protein only.	
CC	The gene and encoded protein may be used in the prevention, treatment and	
CC	diagnosis of diseases associated with inappropriate apoptosis associated	
CC	protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,	
CC	melanomas and lymphomas), inflammation, autoimmune disorders (e.g.	
CC	multiple sclerosis) and viral infections (e.g. herpes).	
XX	Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;	
SQ	Query Match 94.9%; Score 1787.2; DB 3; Length 2045;	
	Best Local Similarity 96.8%; Pred. No. 0;	
	Matches 1858; Conservative 0; Mismatches 3; Indels 59; Gaps 1.	
QY	23 CCAAGAAAGATCAAAACAAAGTGGCTGGGCGCGCTGGCCAGAGATCATTCGACGCCAAGAT 82	
Db	20 CCGGAAAGATCAAAACAAAGTGGCTGGGCGCGCGCTGGCCAGAGATCATTCGACGCCAAGAT 79	
QY	83 CTGGCGCGGGTCTTAGCTGTTGCGCCCTCCCTCCCGGGAATGGCGTATCCGGGTGGA 142	
Db	80 CTGGCGCGGGTCTTAGCTGTTGCGCCCTCCCTCCCGGGAATGGCGTATCCGGGTGGA 139	
QY	143 CCCCAGCGCCCTGCTGGGAGGAGATGAGTGCCTGACTTAAAGGATGCTGTCCTTC 202	
Db	140 CCCCAGCGCCCTGCTGGGAGGAGATGAGTGCCTGACTTAAAGGATGCTGTCCTTC 199	
QY	203 ACCGTATGTTGAGAGTGGTGGGCGGGAATGACCGAGTGCAGCTGGAGCTCTTGACCT 262	
Db	200 ACCGTATGTTGAGAGTGGTGGGCGGGAATGACCGAGTGCAGCTGGAGCTCTTGACCT 259	
QY	263 TTCTGCTGAGTGAAGGCTCTTGGGCCCGCGGAGCTTAAGCCCGGCGCCGAGCGGCTTAA 322	
Db	260 TTCTGCTGAGTGAAGGCTCTTGGGCCCGCGGAGCTTAAGCCCGGCGCCGAGCGGCTTAA 319	

QY	323	AGCTCTCTGAGAGTGGAGCGCCGCGGAGAGTGGGCGAGAGCAACTGCGCGCTGCG	382
Db	320	AGCTCTCTGAGAGTGGAGCGCCGCGGAGAGTGGGCGAGAGCAACTGCGCGCTGCG	379
QY	383	GGCAACTCTGCGCGCTGCTGCGCCGCGCAGACTGCTGCGCACTGGCGCGAGCGGC	442
Db	380	GGCAACTCTGCGCGCTGCTGCGCCGCGCAGACTGCTGCGCACTGGCGCGAGCGGC	439
QY	443	GCCGCGCAGTGTCTCCAAACGCTATACCTATGGAACCTCAGCGCTCTTCAAGAGACAG	502
Db	440	GCCGCGCAGTGTCTCCAAACGCTATACCTATGGAACCTCAGCGCTCTTCAAGAGACAG	499
QY	503	AGGGTAGCTGCGCTGCGCGCTGCGCAGTGAAGCACTTCTGCAAAATCTCAGCAGGCTAGT	562
Db	500	AGGGTAGCTGCGCGCTGCGCGCTGCGCAGTGAAGCACTTCTGCAAAATCTCAGCAGGCTAGT	559
QY	563	GGGAGACAGGCTCCCCCGCCCAACCAAGCGGCAAGCGCGGAGTGTGGGGCGCGCCAGTGGTG	622
Db	560	GGGAGACAGGCTCCCCCGCCCAACCAAGCGGCAAGCGCGGAGTGTGGGGCGCGCCAGTGGTG	619
QY	623	GTGCACAGCGCGCGGAGAGAGGGGCGCCAGCGCGCAACCCAGCAGAGTCAAGAGCCCGCA	682
Db	620	GTGCACAGCGCGCGGAGAGAGGGGCGCCAGCGCGCAACCCAGCAGAGTCAAGAGCCCGCA	679
QY	683	GACCTTCTCTGAAGGCAAGTGAACCTGTGACATCCGACTCCGAGTTTGAAGCAGTACT	742
Db	680	GACCTTCTCTGAAGGCAAGTGAACCTGTGACATCCGACTCCGAGTTTGAAGCAGTACT	739
QY	743	GCGAGCAATGGGCGCGCTTGGAGCAGAGGGGCGTGGCATCCCGCGGCGCCAGAGCGCTGGGCG	802
Db	740	GCGAGCAATGGGCGCGCTTGGAGCAGAGGGGCGTGGCATCCCGCGGCGCCAGAGCGCTGGGCG	799
QY	803	GGCAGCTGGAAGTGTGTGGGAGGCGCACCGCAGTGTCTGCGCTCAAGGAGCACTGGGCTGTG	862
Db	800	GGCAGCTGGAAGTGTGTGGGAGGCGCACCGCAGTGTCTGCGCTCAAGGAGCACTGGGCTGTG	859
QY	863	TGCTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAAGGCTTCTGGGGCGACTACC	922
Db	860	TGCTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAAGGCTTCTGGGGCGACTACC	919
QY	923	TGAGTGGGCGCCCTGTGTGAGGCGCCCTCGGAGGCGTGTCTGTACTAGGAGCCCTGGCAGAGG	982
Db	920	TGAGTGGGCGCCCTGTGTGAGGCGCCCTCGGAGGCGTGTCTGTACTAGGAGCCCTGGCAGAGG	979
QY	983	CTGTGGGCGGGAAGGCTTTCGCTCTGCTGATGAGTGGATGAGGCTGACTATAGAGCTG	1042
Db	980	CTGTGGGCGGGAAGGCTTTCGCTCTGCTGATGAGTGGATGAGGCTGACTATAGAGCTG	1039
QY	1043	GCCGCGCGCGCTGTGTGCTGATGAGAGAGAAAGGGGGGCGCGCCGACAGAGGCTTCT	1102
Db	1040	GCCGCGCGCGCTGTGTGCTGATGAGAGAGAAAGGGGGGCGCGCCGACAGAGGCTTCT	1099
QY	1103	GATCCAGGACTGGCAGAGATTGATCCCACTCCAAATCTTCGGGCGCACTTCTCCTGGAGG	1162
Db	1100	GATCCAGGACTGGCAGAGATTGATCCCACTCCAAATCTTCGGGCGCACTTCTCCTGGAGG	1159
QY	1163	GACGACCATCTCTA-----	1176
Db	1160	GACGACCATCTCTA-----	1176
QY	1177	-----CCCTTGAAGGCGCCCTCCCAAGAGTGGGGCTGAGGGCCTTAAC	1222
Db	1220	CCGGGACAGCAGGCGCCCTTGTGACGCGCTTCCACAGAGTGGGGCTTGTAGGGCCTTAAC	1279
QY	1224	ATTTCAGACTGAGTTTCTTCTCCAGACTCTCTTAACCCCAAGGTGTGCGCCCTTAGCTC	1283
Db	1280	ATTTCAGACTGAGTTTCTTCTCCAGACTCTCTTAACCCCAAGGTGTGCGCCCTTAGCTC	1339
QY	1284	CGGAGGCGGGGCGCTGGGCGCTGATCTTCAAGAGGAGGGGCGACGTTACACTACCAAA	1344
Db	1340	CGGAGGCGGGGCGCTGGGCGCTGATCTTCAAGAGGAGGGGCGACGTTACACTACCAAA	1399

Qy	1344	GGCCCCCTGGACAATTGATCTCGATCTTGGGCTGTCTGCACTGTCAAGGTGCACACA	1403
Dp	1400	GGCCCCCTGGACAATTGATCTCGATCTTGGGCTGTCTGCACTGTCAAGGTGCACACA	1459
Qy	1404	CTGGCTGATGTCACACCTGGCCCTGGCTGAGATCTTCCCTGGGCTCTGGCCCTGGCT	1463
Dp	1460	CTGGCTGATGTCACACCTGGCCCTGGCTGAGATCTTCCCTGGGCTCTGGCCCTGGCT	1519
Qy	1464	TCCGACACACACCTTCTTTGGCCTTAAGGGCTTCTCTCAAGACCTCTTAATTGGACACA	1523
Dp	1520	TCCGACACACACCTTCTTTGGCCTTAAGGGCTTCTCTCAAGACCTCTTAATTGGACACA	1579
Qy	1524	ACCAACTCTGGGCTTCAGCCACATATAGTGGGCACTGAGAGCTGGGGTGCACATGGGGCTTGC	1583
Dp	1580	ACCAACTCTGGGCTTCAGCCACATATAGTGGGCACTGAGAGCTGGGGTGCACATGGGGCTTGC	1639
Qy	1584	TCACCTTGGCCACACATCTCCAGCCAGCCACGAGGCCCTGCCAGCTTCAATTYACAGACT	1643
Dp	1640	TCACCTTGGCCACACATCTCCAGCCAGCCACGAGGCCCTGCCAGCTTCAATTYACAGACT	1699
Qy	1644	GACTCTCCTGACCTTCCCCCCTGTCTGTCCAGAGCTGGAATAGACTTGGCACTTGGATGTC	1703
Dp	1700	GACTCTCCTGACCTTCCCCCCTGTCTGTCCAGAGCTGGAATAGACTTGGCACTTGGATGTC	1759
Qy	1704	ACCTGGAGTGTACATGGGAGTGTATTGCGACATCATCCAAAGGCTTACTGTTCACAT	1763
Dp	1760	ACCTGGAGTGTACATGGGAGTGTATTGCGACATCATCCAAAGGCTTACTGTTCACAT	1819
Qy	1764	GGGGCCAAAACCAAGTAACAGCCACCTTCTTGGAAAAGGAATGAAAAGGCTTTGGGGGTG	1823
Dp	1820	GGGGCCAAAACCAAGTAACAGCCACCTTCTTGGAAAAGGAATGAAAAGGCTTTGGGGGTG	1879
Qy	1824	ATGGAANAAGACTTTTCAATATGATACCAATTAACTGCCCTGGAAAAGGGCATAGTGGG	1883
Dp	1880	ATGGAANAAGACTTTTCAATATGATACCAATTAACTGCCCTGGAAAAGGGCATAGTGGG	1939

XX	AAH33283	
ID	AAH33283	standard; cDNA; 2045 BP.
AC	AAH33283;	
DT	03-SEP-2001	(first entry)
XX		
DE	Human colon cancer antigen encoding cDNA SEQ ID NO:339.	
XX		
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection; colorectal carcinoma; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200122920-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000MO-US026524.	
XX		
PR	29-SEP-1999; 99US-0157137P.	
PR	03-NOV-1999; 99US-0163280P.	
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Barash SC, Birse CE, Rosen CA;	
DR	WPI; 2001-235357/24.	
DR	P-PsDB; AAG73852.	
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides useful for preventing, diagnosing and/or treating colorectal cancers.	
PS	Claim 1; Page 2451; 9803pp; English.	
XX		

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 662 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX

XX Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;

Query Match 94.9%; Score 1787.2; DB 4; Length 2045;

Best Local Similarity 96.8%; Pred. No. 0;
Matches 1858; Conservative 0; Mismatches 3; Indels 59; Gaps 1;

QY 23 CCGAAGAGATCCAAACAAAGTGGCTGGGCGCTGCGCCAGAGTCAATCGAGCCCAAGT 82
DB 20 CCGAAGAGATCCAAACAAAGTGGCTGGGCGCTGCGCCAGAGTCAATCGAGCCCAAGT 79
QY 83 CTGGCGGGGTTCTGAGCTTGTTCGGCTCCCTCCCGGGAGATGGGCTATCCGGGATCA 142
DB 80 CTGGCGGGGTTCTGAGCTTGTTCGGCTCCCTCCCGGGAGATGGGCTATCCGGGATCA 139
QY 143 CCGCGGCGCGCTGCTGGAGAGAGATGAGTGCCTGAGTACTAATCGGAGATGCTGCTTC 202
DB 140 CCGCGGCGCGCTGCTGGAGAGAGATGAGTGCCTGAGTACTAATCGGAGATGCTGCTTC 199
QY 203 ACCGTATGTTGAGGTGTGGGCGGCACTGACCGAGTGCAGTGTGAGACTCTTGAGCT 262
DB 200 ACCGTATGTTGAGGTGTGGGCGGCACTGACCGAGTGCAGTGTGAGACTCTTGAGCT 259
QY 263 TTCTGCTGATGAGGCTCTGGGCGCGCGGAGCTTAAGCCGGGCGCGAGGCGCTAG 322
DB 260 TTCTGCTGATGAGGCTCTGGGCGCGCGGAGCTTAAGCCGGGCGCGAGGCGCTAG 319
QY 323 AGCTCTGTGAGTGAAGTGAAGCGCGCGGAGTGCAGGAGAGCAACTGCGGCTGCTG 382
DB 320 AGCTCTGTGAGTGAAGTGAAGCGCGCGGAGTGCAGGAGAGCAACTGCGGCTGCTG 379
QY 383 GGGCACTCTGCGCGCTGCTGGCGCGCAAGCTGCTGCGCACTGCGCGCACTGCGCG 442
DB 380 GGGCACTCTGCGCGCTGCTGGCGCGCAAGCTGCTGCGCACTGCGCGCACTGCGCG 439
QY 443 GCGGCGCAGTGTCTCCAGAAAGCTATAGTGGCACTCAGCTCTTCAAGAGAGCA 502
DB 440 GCGGCGCAGTGTCTCCAGAAAGCTATAGTGGCACTCAGCTCTTCAAGAGAGCA 499
QY 503 AGGATGCTGCGCTGCGCGCTGCGAGTCAAGCAAGTCTGCAAAATCTCAAGAGAGT 562
DB 500 AGGATGCTGCGCTGCGCGCTGCGAGTCAAGCAAGTCTGCAAAATCTCAAGAGAGT 559
QY 563 GGGAGACAGGCTCCCCCAACCAAGGCGAGCGGAGTCCGGGCGCGGCGAGTGG 622
DB 560 GGGAGACAGGCTCCCCCAACCAAGGCGAGCGGAGTCCGGGCGCGGCGAGTGG 619
QY 623 GTCGCAAGCGGCGGAGGAGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 682
DB 620 GTCGCAAGCGGCGGAGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 679
QY 683 GACCTTCTCTGAGAGGCAAGTGAAGTGAATCCGAGTCCGGGTTCCAGAGAGTACT 742
DB 680 GACCTTCTCTGAGAGGCAAGTGAAGTGAATCCGAGTCCGGGTTCCAGAGAGTACT 739

QY 743 GGGAGATGAGGCGAGCTTTGAGAGAGGAGGAGTGGCAATCCCGGCGGCGGCGGCTGGCGC 802
DB 740 GGGAGATGAGGCGAGCTTTGAGAGAGGAGGAGTGGCAATCCCGGCGGCGGCGGCTGGCGC 799
QY 803 GCGAGCTGAGACTGTTTGGGCGAGGCGGAGGAGTGGCTGCTCAAGAGGAGCTGGGCTG 862
DB 800 GCGAGCTGAGACTGTTTGGGCGAGGCGGAGGAGTGGCTGCTCAAGAGGAGCTGGGCTG 859
QY 863 TGGTTTGTGACATCAAGTTCAGAGCTCTCTATCTGAGAGCGCTTCTGGGCGAGTAC 922
DB 860 TGGTTTGTGACATCAAGTTCAGAGCTCTCTATCTGAGAGCGCTTCTGGGCGAGTAC 919
QY 923 TGAAGTGGCGGCGGCTGAGAGGCGGCGGAGGAGTGGCTGAGAGGCGGCGGAGGAG 982
DB 920 TGAAGTGGCGGCGGCTGAGAGGCGGCGGAGGAGTGGCTGAGAGGCGGCGGAGGAG 979
QY 983 CTGTGGGCGGAGGAGCTGTTCCGCTGCTGATGAGTGGATGAGCTGACTATGAGAGCTG 1042
DB 980 CTGTGGGCGGAGGAGCTGTTCCGCTGCTGATGAGTGGATGAGCTGACTATGAGAGCTG 1039
QY 1043 GCGGCGGCGGCTGTTGCTGATGAGAGAGAGAGGAGGAGGCGGCGGCGGAGAGGCTCT 1102
DB 1040 GCGGCGGCGGCTGTTGCTGATGAGAGAGAGAGGAGGAGGCGGCGGCGGAGAGGCTCT 1099
QY 1103 GATCCAGAGCTGAGAGAGATGATGATCCCACTCCAGATCCCGGCGGAGAGCTTCT 1162
DB 1100 GATCCAGAGCTGAGAGAGATGATGATCCCACTCCAGATCCCGGCGGAGAGCTTCT 1159
QY 1163 GACGACATCTCTA----- 1176
DB 1160 GACGACATCTCTA----- 1219
QY 1177 -----CCCTTGAAGCGGCTCCAGAGATGAGGCTTGAAGCTTAAAC 1223
DB 1220 CCGGAGCAGAGGCGGCTTGAAGGCGGCTCCAGAGATGAGGCTTGAAGCTTAAAC 1279
QY 1224 ATTTCCAGTGAATTCCTTCCAGAGCTCCCTGAGAGGAGGAGGAGGAGGAGGAGGAG 1283
DB 1280 ATTTCCAGTGAATTCCTTCCAGAGCTCCCTGAGAGGAGGAGGAGGAGGAGGAGGAG 1339
QY 1284 CCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1343
DB 1340 CCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1399
QY 1344 GGGCGGCTGCAATTTGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1403
DB 1400 GGGCGGCTGCAATTTGATCTTGAATCTTGAATCTTGAATCTTGAATCTTGAATCTTGA 1459
QY 1404 CTGCTGATGCTCAAGCGGCGGCTGAGAGTCTTCCGAGGAGGAGGAGGAGGAGGAGGAG 1463
DB 1460 CTGCTGATGCTCAAGCGGCGGCTGAGAGTCTTCCGAGGAGGAGGAGGAGGAGGAGGAG 1519
QY 1464 TCCAGAGCAGACTTCTTGGGCTTGAAGGCTTCTCTCAAGAGGAGGAGGAGGAGGAGGAG 1523
DB 1520 TCCAGAGCAGACTTCTTGGGCTTGAAGGCTTCTCTCAAGAGGAGGAGGAGGAGGAGGAG 1579
QY 1524 ACCAAGCTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1583
DB 1580 ACCAAGCTGGGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1639
QY 1584 TCACTTGGGCGAGACTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1643
DB 1640 TCACTTGGGCGAGACTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1699
QY 1644 GACTCTCTCACTTCCCGGCTGCTGCTCAAGAGTGAATCAATGAGCTTGAATGAGTGC 1703
DB 1700 GACTCTCTCACTTCCCGGCTGCTGCTCAAGAGTGAATCAATGAGCTTGAATGAGTGC 1759
QY 1704 ACTGAGAGTGAATGAGGAGGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1763
DB 1760 ACTGAGAGTGAATGAGGAGGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1819
QY 1764 GGGGCGCAAAACAGTAAAGCAGCACTTCTTGGAGAGGAGATGCAAGGCTTGGGCGG 1823

Db 1820 GGGGCCCCAAACAGTAACAGCCACTTGTGAAAGGAATGCAAGGCTTTGGGGGTG 1879
Qy 1824 ATGAAAGAGCTTTTCAATATGATACCAATTAACTGAGGGAAGGATGAGG 1883
Db 1880 ATGAAAGAGCTTTTCAATATGATACCAATTAACTGAGGGAAGGATGAGG 1939

RESULT 5
AAFI8296
ID AAFI8296 standard; DNA, 2044 BP.
XX AAFI8296;
AC
XX
XX 14-MAR-2001 (first entry)
DE Lung cancer associated polynucleotide sequence SEQ ID 315.
XX
XX Human; lung cancer associated protein; neuroprotective; cytoskeletal;
XX cardioprotective; immunomodulatory; muscular active; vulnerable;
XX gastrointestinal; nephrotoxic; antiinfective; gynecological;
XX antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX proliferative disorder; wound healing; infectious disease; ds.
XX Homo sapiens.
XX WO200055180-A2.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005918.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX
XX Ruben SM;
XX
XX WPI; 2000-587514/55.
XX P-PSDB; AAB58420.
XX
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
XX PT antigens, useful for treatment, prevention, and diagnosis of disorders
XX PT such as lung cancer.
XX
XX Claim 1, Page 772-773, 1425pp; English.
XX
XX Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer
XX CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
XX CC associated proteins and polynucleotide sequences, their agonists, and
XX CC antagonists may have neuroprotective; cytoskeletal; cardioprotective;
XX CC immunomodulatory; muscular active general; vulnerable; gastrointestinal
XX CC general; nephrotoxic; antiinfective; gynecological; or antibacterial
XX CC activity. The invention also includes antibodies specific for the protein
XX CC or polynucleotide sequences. The lung cancer associated polynucleotide
XX CC sequences may be used for detection of lung cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The proteins may be used to treat disorders such as
XX CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
XX CC cardiovascular, renal, and proliferative disorders. The proteins may also
XX CC be used in the treatment of wounds and infectious diseases.
XX CC Polynucleotide sequences AAFI8425 - AAFI8433 and peptide AAB58549 are
XX CC used in the course of the invention for the identification and
XX CC characterisation of the polynucleotide and protein sequences
XX
XX Sequence 2044 BP; 457 A; 620 C; 603 G; 364 T; 0 U; 0 Other;

Query March 94.3%; Score 1775.2; DB 3; Length 2044;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1857; Conservative 0; Mismatches 3; Indels 60; Gaps 2;
Qy 23 CCAAGAAAGATCCAAACAGTGTGCTGGCGCCGCTGCCAGAGATCATCGACGCCAGAAAT 82

Db 20 CCGAAAGATCCAAACAGTGTGCTGGCGCCGCTGCCAGAGATCATCGACGCCAGAAAT 79
Qy 83 CTGGCCGGGCTTCTGAGCTTGTTCGGCTCCCTCCCGGGGAATGGGCTATCGGGCTCA 142
Db 80 CTGGCCGGGCTTCTGAGCTTGTTCGGCTCCCTCCCGGGGAATGGGCTATCGGGCTCA 139
Qy 143 CCGCGGCCCGCTGCTGGAGAGAGATGATGCTGAGCTTACTAGGGATCTCTGCTTC 202
Db 140 CCGCGGCCCGCTGCTGGAGAGAGATGATGCTGAGCTTACTAGGGATCTCTGCTTC 199
Qy 203 ACCGTAATGTTGAGGTGTGGGCGGGCACTGACCGAGTGCAGCTTGAGACTCTTGACT 262
Db 200 ACCGTAATGTTGAGGTGTGGGCGGGCACTGACCGAGTGCAGCTTGAGACTCTTGACT 259
Qy 263 TTTCTGCTGAGATGAGGCTTCTGGGCGCGCGGAGGCTTAAAGCCGGGCGCGAGGGCTAG 322
Db 260 TTTCTGCTGAGATGAGGCTTCTGGGCGCGCGGAGGCTTAAAGCCGGGCGCGAGGGCTAG 319
Qy 323 AGCTCTGCTGAGAGCTGAGACCGCGGCGGAGTGCAGGAGCAACTGCTGGCTGCTGG 382
Db 320 AGCTCTGCTGAGAGCTGAGACCGCGGCGGAGTGCAGGAGCAACTGCTGGCTGCTGG 379
Qy 383 GGCACCTCTGCGCGTGTGCTGGCCCGCAGCACTGTGCTGCGCACTTGGCGGCAAGCGGC 442
Db 380 GGCACCTCTGCGCGTGTGCTGGCCCGCAGCACTGTGCTGCGCACTTGGCGGCAAGCGGC 439
Qy 443 GCGCGGCGAGGTGTCTCCAGAACGCTATAGCTATGAGCACTTCACTCTTCAAAAGAGACAG 502
Db 440 GCGCGGCGAGGTGTCTCCAGAACGCTATAGCTATGAGCACTTCACTCTTCAAAAGAGACAG 499
Qy 503 AGGTAAGCTGCGCTGCGCGCTGCGCAGTCAAGCAAGTCTGCAAAATTCACAGAGGTCAGT 562
Db 500 AGGTAAGCTGCGCTGCGCGCTGCGCAGTCAAGCAAGTCTGCAAAATTCACAGAGGTCAGT 559
Qy 563 GGGAGACAGGCTTCCCCCAGCAAGCGGCGAGGCGGAGTGGGCGGCGGCGGCGGCGGCGG 622
Db 560 GGGAGACAGGCTTCCCCCAGCAAGCGGCGAGGCGGAGTGGGCGGCGGCGGCGGCGGCGG 619
Qy 623 GTGTCAGACGCGCGCGGAGAGAGGGGCCCCAGCCGCACTCCAGAGAGCATGACAGCCGCGCA 682
Db 620 GTGTCAGACGCGCGCGGAGAGAGGGGCCCCAGCCGCACTCCAGAGAGCATGACAGCCGCGCA 679
Qy 683 GACCTTCTCTGAAAGGCAAAAGTACCTGTGATCATTCGCGCTCGGGGTTCCAGCAGAGTACT 742
Db 680 GACCTTCTCTGAAAGGCAAAAGTACCTGTGATCATTCGCGCTCGGGGTTCCAGCAGAGTACT 739
Qy 743 GCGAGCATGGGCGAGCCTTGGAGCAGGGCGTGGCATTCGCGCGGCGCCAGAGGCGTGGCGC 802
Db 740 GCGAGCATGGGCGAGCCTTGGAGCAGGGCGTGGCATTCGCGCGGCGCCAGAGGCGTGGCGC 799
Qy 803 GCGAGCTGAGACGTGTTTGGGCGAGGCCACCGCATGTCTGCGCTCAAGAGACTTGGGCTCTG 862
Db 800 GCGAGCTGAGACGTGTTTGGGCGAGGCCACCGCATGTCTGCGCTCAAGAGACTTGGGCTCTG 859
Qy 863 TGGTTTGTGATCAATCAAGTTTCTGAGGCTCTCATCTGAGCGGCTTCTGGGGGCGACTACC 922
Db 860 TGGTTTGTGATCAATCAAGTTTCTGAGGCTCTCATCTGAGCGGCTTCTGGGGGCGACTACC 919
Qy 923 TGAATGGCGGCTTCTGTCAGGCGCTTGGCGGCGGTGTTCTGACTGAGGCGCTTGGAGAGG 982
Db 920 TGAATGGCGGCTTCTGTCAGGCGCTTGGCGGCGGTGTTCTGACTGAGGCGCTTGGAGAGG 978
Qy 983 CTGTGGGCGGAGAGGCTGTTGCGCTGTGTGATGATGATGAGGCTGACTATGAGGCTG 1042
Db 979 CTGTGGGCGGAGAGGCTGTTGCGCTGTGTGATGATGATGAGGCTGACTATGAGGCTG 1038
Qy 1043 GCGCGGCGGCGCTGTTGTGATGATGAGAGAGAGAGGAGGCGGCGCGGAGCAGAGGCTCTCT 1102
Db 1039 GCGCGGCGGCGCTGTTGTGATGATGAGAGAGAGAGGAGGCGGCGCGGAGCAGAGGCTCTCT 1098
Qy 1103 GATTCAGAGCTGGCAGAGATTTGATCCCACTCAAGTCTTCGGGCGCACTTCTCTGAGAG 1162

[illegible]

Oy		1177	--CCCTTGAAGAGCCCTCCACAGGATGTGGAGCTGTAGAGGCCTAAACCATTTCACACT	1233
Dd		1200	AAGGCCCTTTAAGACCCCCTCCACAGGATGTGGAGCTGTAGAGGCCTAAACAATTTCACACT	1259
Oy		1234	GAGTTTCTTTCOCAGACTCCTCCTAACCCAGGTGTGCCCCCTTAAGCTCCGAGAGCGGG	1293
Dd		1260	GAGTTTCTTTCOCAGACTCCTCCTTA--CCCAAGGTGTCCCATATGGCTCCGAGACGGGG	1318
Oy		1294	GGCTGGGCGCTGTATCTCAGAAGGGAGGGGACACACTACACATCACCAAAGGCCCTCTG	1353
Dd		1319	GGCTGGGCGCTGTATCTCAGAAGGGAGGGGACACACTACACCAAAGGCCCTCTG	1378
Oy		1354	CACATTGTATCTCGATCTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1413
Dd		1379	CACATTGTATCTCGATCTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1438
Oy		1414	CTCACACTGGCCCCCTGTAGATCTTCCCTGGGCTGTGCCCTGGCTGTCTTCCAGACACA	1473
Dd		1439	CTCACACTGGCCCCCTGTAGATCTT--CTTGGGCTGTGCCCTGGCTGT--TTCCAGACACA	1496
Oy		1474	CACTTCTTTGGCCTTAGGGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1530
Dd		1497	CACTTCTTTGGCCTTAGGGCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1556
Oy		1531	TGGGCTTTCAGCC--ACATCAGTGTGGGCACTGGAGCTGGGTGTGTGTGTGTGTGTGTGTGTGT	1589
Dd		1557	GGGATTTTCAGCCAAATCAGTGTGGGCACTGGAGCTGGGTGTGTGTGTGTGTGTGTGTGTGT	1616
Oy		1590	TGCCACACATCTTCCAGCCAGCCAGGGCCCTGCCAGCTTCAATTTCACAGCTGACTCT	1649
Dd		1617	TGCCACAAATTTTCCAGCCAGCCAGGGCCCTGCCAGCTTCAATTTCACAGCTGACTCTNT	1676
Oy		1650	CCTCACCTTCCCCCTGT	1709
Dd		1677	CCTCACCTTCCCCCTGT	1736
Oy		1710	AGTGTACATGAGGAGTGTATGTGCAGATATATACCAAGGCTTACTGTGTGTGTGTGTGTGTGT	1759
Dd		1737	AGTGTACATGAGGAGTGTATGTGCAGATATATACCAAGGCTTACTGTGTGTGTGTGTGTGTGT	1796
Oy		1770	AAAACGAGTAAACAGCCACCTTCTGTGAAAGGGAAATGCAAAAGGCTTGGGGGGTGTATGAA	1829
Dd		1797	AAAACGAGTAAACAGCCACCTTITTTGAAAGGGAAATGCAAAAGGCTTGGGGGGTGTATGAA	1856
Oy		1830	AAGACTTTTATCAATGATATCAATTAACTGGCCTGTGAAAGGGCATAGGTGGG	1883
Dd		1857	AAGACTTTTATCAATGATATCAATTAACTGGCCTGTGAAAGGGCATAGGTGGG	1910
<hr/>				
RESULT 7				
AAD59062 standard; cDNA; 1924 BP.				
XX	AAD59062;			
AC	AAD59062;			
XX	18-DEC-2003 (first entry)			
DT	Human DED4 full length gene.			
XX				
DE				
KM	Human, death Domain; DD; death effector domain; DED; cell proliferation;			
KW	Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;			
KW	neural growth factor receptor-interacting death domain; cell adhesion;			
KW	vastrotopic; microbial infection; inflammation; allograft rejection; CTDD;			
KW	cell stress response; benign prostatic hyper trophy; antibacterial; NTD;			
KM	apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;			
KM	neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;			
KM	keloid; gene; ss.			
XX				
OS	Homo sapiens.			
XX				
FT	Key	Location/Qualifiers		
CD	CDS	91..1044		
LT		/+tag= a		

FT /product= "Human DED4 protein"
 FT /note= "No stop codon"
 FT /partial
 FT misc_signal 157..222
 FT /tag= b
 FT /note= "Nuclear localisation sequence"
 XX US2003049702-A1.
 PD 13-MAR-2003.
 PF 15-NOV-2001; 2001US-00001254.
 XX
 PR 17-NOV-2000; 2000US-00715893.
 PR 17-NOV-2000; 2000US-0367360P.
 PR 29-JUN-2001; 2001US-0301889P.
 XX
 PA (REED/) REED J C.
 PA (GODZIK/) GODZIK A.
 PA (PAWL/) PAWLOWSKI K.
 PA (FIOR/) FIORENTINO L.
 PA (LEES/) LEE S H.
 PA (ROTH/) ROTH W.
 PA (STEN/) STENNER-LIEWEN F.
 XX
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
 PI Stenner-Liewen F,
 XX
 DR MPI; 2002-500222/53.
 DR P-Psdb; AAE38903.
 XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
 XX
 PS Claim 19; Page 42-44; 99pp; English.
 XX
 CC The present invention provides novel death Domain (DD) and death effector
 CC domain (DED) proteins and nucleic acids encoding them. The invention also
 CC provides death domain containing protein such as Chlamydia trachomatis
 CC death domain containing protein (CTPD) DD and neutral growth factor
 CC receptor-interacting death domain (NIDD) DD. The invention is useful for
 CC identifying a binding agent (e.g. protein or drug) that binds a DD, DED
 CC or NB-ARC domain from DAP3, IRAK4, CTRD, DED4 or NIDD with a candidate
 CC binding agent and identifying an effective agent (e.g. protein or drug)
 CC that modulates the association of a DD, DED or NB-ARC domain with protein
 CC that binds the DD, DED or NB-ARC domain. The invention is also useful for
 CC modulating the level of cell process such as apoptosis, cell adhesion,
 CC cell proliferation, cell stress responses, responses to microbial
 CC infection and B cell immunoglobulin class switching. Dds, DEDs and NB-ARC
 CC domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are
 CC useful for discovery of drugs that suppress infection, autoimmunity,
 CC inflammation, allergy, allograft rejection, sepsis and other diseases.
 CC
 CC DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
 CC autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
 CC hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
 CC inflammatory hyperplasia and smooth muscle cell proliferation in arteries
 CC following balloon angioplasty (restenosis). The invention is also used in
 CC antibody therapy and gene therapy. The present sequence is human DED4
 CC full length gene. The DED4 gene is located on chromosome 19
 XX
 SQ Sequence 1924 BP; 360 A; 599 C; 594 G; 368 T; 0 U; 3 Other:

Query Match 89.3%; Score 1681.4; DB 6; Length 1924;
 Best Local Similarity 95.5%; Pred. No. 0;
 Matches 1828; Conservative 0; Mismatches 19; Indels 67; Gaps 7;

QY 33 CCAAAACAACTGCTGCGCGCTGCGCCAGAGATCATTCGAGCCGAGATCTGCGCGGT 92
 Db 1 CCAAAACAACTGCTGCGCG-CTGCGCCAGAGATCATTCGAGCCGAGATCTGCGCGGT 59
 QY 93 TCTGAGCTTGTTCGCTCCCTCCCGGAGATGGGCTATCCGGGTCCAGCCCGGCGCC 152

Db 60 TCTGAGCTTGTTCGCTCCCTCCCGGAGATGGGCTATCCGGGTCCAGCCCGGCGCC 119
 QY 153 GTGCTGGAGAGAGATGAGTGTCTGAGTACTTACGAGATCTGTGCTTACCGTATGTT 212
 Db 120 GTGCTGGAGAGAGATGAGTGTCTGAGTACTTACGAGATCTGTGCTTACCGTATGTT 179
 QY 213 CGAGGTGGGGCGGCGCACTGACCGAGTGGAGTGGAGTCTTCGGGCTTCTGTGTGA 272
 Db 180 CGAGGTGGGGCGGCGCACTGACCGAGTGGAGTGGAGTCTTCGGGCTTCTGTGTGA 239
 QY 273 TGAGGCTCTGCGCGCGCGAGGCTTATGCTCGGGCCGCGAGCGGCTTATGAGCTCTGTG 332
 Db 240 TGAGGCTCTGCGCGCGCGAGGCTTATGCTCGGGCCGCGAGCGGCTTATGAGCTCTGTG 299
 QY 333 GAGCTGAGCGCGCGGCGCACTGCGCGAGAGCAACTGCGGCTGTGGGCAACTCTT 392
 Db 300 GAGCTGAGCGCGCGGCGCACTGCGCGAGAGCAACTGCGGCTGTGGGCAACTCTCT 359
 QY 393 GCGGCTGCTGGCGCGCGCGAGCTGTGCGGCACTTGGCGCGCAAGCGGCGCGCGAGT 452
 Db 360 GCGGCTGCTGGCGCGCGCGAGCTGTGCGGCACTTGGCGCGCAAGCGGCGCGAGT 419
 QY 453 GTTCCAGAACGCTATGCTATGCACTTCAAGTCTTCAAGAGAGACAGAGGTAAGTGTG 512
 Db 420 GTTCCAGAACGCTATGCTATGCACTTCAAGTCTTCAAGAGAGACAGAGGTAAGTGTG 479
 QY 513 CCGTGGCGCTGCGGAGTCAAGAGATTCTGCAAAATTTCTGACGAGGTCAGTGGAGACAG 572
 Db 480 CCGTGGCGCTGCGGAGTCAAGAGATTCTGCAAAATTTCTGACGAGGTCAGTGGAGACAG 539
 QY 573 CTCGCCCGCAACAGCGCGAGCGGCGAGTGGGGCGCGGCGCAAGTGGTGGCAGCG 632
 Db 540 CTCGCCCGCAACAGCGCGAGCGGCGAGTGGGGCGCGGCGCAAGTGGTGGCAGCG 599
 QY 633 GCGGCGAGAGAGGCGCGCGAGCGGCGAGTGGGGCGCGGCGCAAGTGGTGGCAGCG 692
 Db 600 GCGGCGAGAGAGGCGCGCGAGCGGCGAGTGGGGCGCGGCGCAAGTGGTGGCAGCG 659
 QY 693 TGAAGCAAAAGTACCTGTGATCTCGGCTCGGGTTGAGGAGATCTGCGAGCATG 752
 Db 660 TGAAGCAAAAGTACCTGTGATCTCGGCTCGGGTTGAGGAGATCTGCGAGCATG 719
 QY 753 GCGAGCTTGGAGAGAGGCGGTCATCCCGGCGCGCGCGCGCGGCGGCGGCACTGGA 812
 Db 720 GCGAGCTTGGAGAGAGGCGGTCATCCCGGCGCGCGCGCGCGGCGGCGGCACTGGA 779
 QY 813 CGTGTGGGCGAGGCGCGAGCTGCGCTCAAGAGGAGCTGCGGCTGTGTTTGTGA 872
 Db 780 CGTGTGGGCGAGGCGCGAGCTGCGCTCAAGAGGAGCTGCGGCTGTGTTTGTGA 839
 QY 873 CATCAAGTTCTGAGAGCTCTCTATCTGAGCGCTTCTGGGGCGCACTACCTGATGGCGC 932
 Db 840 CATCAAGTTCTGAGAGCTCTCTATCTGAGCGCTTCTGGGGCGCACTACCTGATGGCGC 899
 QY 933 CCGTGTGAGGCGCGCTGCGGGGCTGTCTGATCTGAGCGCTTCTGGGGCGGCGG 992
 Db 900 CCGTGTGAGGCGCGCTGCGGGGCTGTCTGATCTGAGCGCGCTTCTGGGGCGGCGG 959
 QY 993 GGAGGCTGTTCGCTGCTGTGATGTGATGAGCTGATGAGGCTGTGAGGCTGTGCGGCGCG 1052
 Db 960 GGAGGCTGTTCGCTGCTGTGATGTGATGAGCTGATGAGGCTGTGAGGCTGTGCGGCGCG 1019
 QY 1053 CCTGTGCTGATGAGAGAGAGAGGCGGCGCGCGAGAGGCGCTCTGATCCAGAGAC 1112
 Db 1020 CCTGTGCTGATGAGAGAGAGAGGCGGCGCGCGAGAGGCGCTCTGATCCAGAGAC 1079
 QY 1113 TGGCAGATTTGATCCCACTCCAAAGTCTCCGGGCGCACTTCTCTGGAGAGACAGCAATC 1172
 Db 1080 TGGCAGATTTGATCCCACTCCAAAGTCTCCGGGCGCACTTCTCTGGAGAGAGCAATC 1139
 QY 1173 TCTA----- 1176
 Db 1140 TCTA----- 1199

OY	1177	----	CCCTTGA	CAGCC	CCCTCC	CA	CAGAT	GTG	GGCT	CTG	AGG	CCCTAA	CA	ATTTC	CA	AGT	1233															
Db	1200	AG	CCCTT	GAC	AGCC	CCCTCC	CA	CAGAT	GTG	GGCT	CTG	AGG	CCCTAA	CA	ATTTC	CA	AGT	1259														
OY	1234	GAG	TTTCT	CCCTG	AC	CTCT	CTTA	CCCC	CAG	GTG	TG	CCCC	CTTA	AG	CTCC	GAG	GGGG	1233														
Db	1260	GAG	TTTCT	CCCTG	AC	CTCT	CTTA	CCCC	CAG	GTG	TG	CCCC	CTTA	AG	CTCC	GAG	GGGG	1318														
OY	1284	GG	CTGGG	GC	TGAT	CT	CA	GA	AGG	GGG	GGG	CA	GT	TA	CA	CA	TA	CA	AA	GGCC	CCCTG	1353										
Db	1319	GG	CTGGG	GC	TGAT	CT	CA	GA	AGG	GGG	GGG	CA	GT	TA	CA	CA	TA	CA	AA	GGCC	CCCTG	1378										
OY	1354	CAC	ATTGAT	CT	CTG	AT	CTT	GGG	CGT	CTG	CA	ACT	GT	CA	CAG	GTG	CA	CA	CT	CG	CT	CA	ATG	1413								
Db	1379	CAC	ATTGAT	CT	CTG	AT	CTT	GGG	CGT	CTG	CA	ACT	GT	CA	CAG	GTG	CA	CA	CT	CG	CT	CA	ATG	1438								
OY	1414	CT	CAC	ACT	GG	CCCT	CTG	TA	GA	TCTT	CC	CTG	GGG	CT	CTG	CC	CTG	GG	CT	CT	CT	CC	CA	GCA	1473							
Db	1439	CT	CAC	ACT	GG	CCCT	CTG	TA	GA	TCTT	CC	CTG	GGG	CT	CTG	CC	CTG	GG	CT	CT	CT	CC	CA	GCA	1496							
OY	1474	CAC	TTCT	TG	GC	CTA	AGG	GGCTT	CT	CT	CT	CA	GGA	CC	CTTA	ATT	GA	CC	CA	CA	---	CCA	ACC	1530								
Db	1497	CAC	TTTCT	TG	GC	CTA	AGG	GGCTT	CTT	CT	CT	CA	GGA	CC	CTTA	ATT	GA	CC	CA	CA	CCA	AA	ACT	1556								
OY	1531	TGG	GCTT	CA	GC	---	ACAT	AG	TGG	GC	CA	CTG	GA	CTG	GA	CTG	GA	CA	TG	GGG	CT	G	CT	CA	CTT	1589						
Db	1557	GGG	GTTT	CA	GC	CA	AA	AT	CA	TGG	GC	CA	CTG	GA	CTG	GA	CA	TG	GGG	CT	G	CT	CA	CTT	1616							
OY	1590	TG	CC	CA	CA	AT	CT	CC	AG	CAG	CA	GA	GGG	CC	TG	CC	AG	CTT	CA	TTA	ATA	CA	CA	CT	GA	CT	1649					
Db	1617	TG	CC	CA	CA	AT	CT	CC	AG	CAG	CA	GA	GGG	CC	TG	CC	AG	CTT	CA	TTA	ATA	CA	CA	CT	GA	CT	1676					
OY	1650	CCT	CA	CTT	CCCC	CC	CTG	CT	GC	AT	CC	AG	CT	GA	CA	T	GA	CTT	GA	CTT	GA	AT	GA	CT	GA	CTT	1709					
Db	1677	CCT	CA	CTT	CCCC	CC	CTG	CT	GC	AT	CC	AG	CT	GA	CA	T	GA	CTT	GA	CTT	GA	AT	GA	CT	GA	CTT	1736					
OY	1710	AG	TG	CA	CAT	GGG	AG	TGT	TA	TG	CA	GA	CA	T	CA	TA	CA	AA	G	CC	TA	CA	T	TG	CA	CA	TG	GG	CC	1769		
Db	1737	AG	TG	CA	CAT	GGG	AG	TGT	TA	TG	CA	GA	CA	T	CA	TA	CA	AA	G	CC	TA	CA	T	TG	CA	CA	TG	GG	CC	1796		
OY	1770	AAA	CC	CA	TAA	A	CA	G	CA	CTT	CTT	TG	GA	AA	G	GA	TG	CA	AA	A	G	CTT	TG	GG	GG	T	GA	TG	GA	1829		
Db	1797	AAA	CC	CA	TAA	A	CA	G	CA	CTT	CTT	TG	GA	AA	G	GA	TG	CA	AA	A	G	CTT	TG	GG	GG	T	GA	TG	GA	1856		
OY	1830	AA	GA	CTT	T	T	CA	AA	T	GA	T	GA	CA	AA	T	T	AA	CA	T	GA	CC	CT	TG	GA	AA	G	GG	CA	T	GA	GG	1883
Db	1857	AA	GA	CTT	T	T	CA	AA	T	GA	T	GA	CA	AA	T	T	AA	CA	T	GA												

RESULT 8	
AA562603	
ID	AA562603 standard; cDNA; 1966 BP.
XX	
AC	AA562603;
XX	
DT	14-FEB-2002 (first entry)
XX	
DE	cDNA sequence #390 encoding novel human secreted protein.
XX	
KM	Human secreted protein; hyperproliferative disorder; autoimmune disorder
KM	immune deficiency disorder; blood disorder; inflammatory disorder;
KM	infectious disorder; gene therapy; antimicrobial; hepatotropic;
KM	immunosuppressive; antirheumatic; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200177291-A2.
XX	
PD	18-OCT-2001.
XX	
PF	29-MAR-2001; 2001WO-US010485.
XX	

PR 06-APR-2000; 2000US-0195604P.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ,
PI Gulukota K, Graham JR;
XX
XX WPI; 2002-010900/01.
XX
XX
XX New polynucleotides encoding secreted proteins useful for treating e.g
PT asthma, HIV and Crohn's disease.
XX
XX
XX Claim 1; Page 280; 391pp; English.

CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The
CC proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are useful
CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
CC immune deficiency disorders (e.g. severe combined immunodeficiency
CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis) and
CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
CC the invention are also useful in gene therapy. AAs62214-AAs62838
CC represent the cDNA sequences of the invention that encode for novel human
CC secreted proteins
XX
XX Sequence 1966 BP; 356 A; 618 C; 614 G; 378 T; 0 U; 0 Other;

Query Match	86.6%	Score 1630.4	DB 6	Length 1966
Best Local Similarly	92.0%	Pred. No. 0		
Matches 1808; Conservative	0	Mismatches	1	Indels 157; Gaps 2

Oy	75	GCACAAATCTGGCCGGGGTTCTGAAGTTCTTCGCTCCCTCCCGGGAAATGGCGCATC	134
Db	1	GCCAGAAATCTGGCCGGGGTTCTGAAGTTCTTCGCTCCCTCCCGGGAAATGGCGCATC	60
Oy	135	CGGGTCGACCCCGGGCCCGGTGCTGGAGAGAGATGAGTGCCTGAGACTTACCGGATGCT	194
Db	61	CGGGTCGACCCCGGGCCCGGTGCTGGAGAGAGATGAGTGCCTGAGACTTACCGGATGCT	120
Oy	195	GTGGCTTCAACCGTAATGTTCTGAAGTGTGGCGGGCACTGACCGAATGCCAGCTGAGACT	254
Db	121	GTGGCTTCAACCGTAATGTTCTGAAGTGTGGCGGGCACTGACCGAATGCCAGCTGAGACT	180
Oy	255	CCTGGCCTTTCGTGCGATGAGGCTCTCGAGCGCGCCGAGAGGCTTAACTCCCGGGCCGAG	314
Db	181	CTGGGCTTTCTGCTGGANAGAGCTCTGGGGCGCCGAGAGGCTTAACTCCCGGGCCGAG	240
Oy	315	CGGCTTAAGAGCTCTGTCTGAGCTTGAAGCCCGCGGCGATGTGGCGAGAGCAACTGTGG	374
Db	241	CGGCTTAAGAGCTCTGTCTGAGAGCTTGAAGCCCGCGGCGATGTGGCGAGAGCAACTGTGG	300
Oy	375	GCCTGTGGAGGCAACTCTGTGCGGTCTGTGGCCCGGCAACGACTGTGCGGACACTGTGGCGG	434
Db	301	GCCTGTGGAGGCAACTCTGTGCGGTCTGTGGCCCGGCAACGACTGTGCGGACACTGTGGCGG	360
Oy	435	CAAGCGGCGCCGGCCAGATGTCTCCAGAAAGCTAATGCTAATGGCACTTCAGCTTTCAAA	494
Db	361	CAAGCGGCGCCGGCCAGATGTCTCCAGAAAGCTAATGCTAATGGCACTTCAGCTTTCAAA	420
Oy	495	GAGGACAGAGGATAGCTGCGCTCGCGCTCGGCAAGTCAAGCAAGTTCTCAAAATTCAGCA	554
Db	421	GAGGACAGAGGATAGCTGCGCTCGCGCTCGGCAAGTCAAGCAAGTTCTCAAAATTCAGCA	480
Oy	555	GGGTCATGTGGAGACAGAGCTTCCCCCAACCAAGCGGAGGCGTGAATGTGGGGCCGGCC	614
Db	481	GGGTCATGTGGAGACAGAGCTTCCCCCAACCAAGCGGAGGCGTGAATGTGGGGCCGGCC	540

QY 615 CAGTGTGTGTCAGACGCGGCGGAGAGGGGCCCCAGCCGACCCAGACGATCA 674
 DB 541 CAGTGTGTGTGTCAGACGCGGCGGAGAGGGGCCCCAGCCGACCCAGACGATCA 600
 QY 675 GCGCGCAGACCTTCTCTGAAGGCAAGTGAACCTG----- 710
 DB 601 GCGCGCAGACCTTCTCTGAAGGCAAGTGAACCTG----- 710
 QY 711 ----- 710
 DB 661 AGCATCTGCTTCTGTGAGGACCTCCGGAAGCTTCATTTCATGTGGAAGGCCAAGGGA 720
 QY 711 -----TGAATCCGGCTCCGGCTTCGAGCAGAGTATCTGCCAGATGAGGCCA 756
 DB 721 GCAAGCTTGTACATATACATCCGGCTTCGGGTTGAGCAGAGTATCTGCCAGATGAGGCCA 780
 QY 757 GCCTTGAGAGAGGCGGTGATGATCCGGCGGCCCCAGGCGCTGGCGGCGGAGCTGAGCGTG 816
 DB 781 GCCTTGAGAGAGGCGGTGATGATCCGGCGGCCCCAGGCGCTGGCGGCGGAGCTGAGCGTG 840
 QY 817 TTGAGGAGGCGCACCGCAGTCTGCGCTCAAGGACCTGGGCTGTGTGTTTGTGACATC 876
 DB 841 TTGAGGAGGCGCACCGCAGTCTGCGCTCAAGGACCTGGGCTGTGTGTTTGTGACATC 900
 QY 877 AAGTTCTCAGAGCTTCTATCTGAGCGCTTCTGGGCGCATCTAGTGAAGTGGCGCCCTG 936
 DB 901 AAGTTCTCAGAGCTTCTATCTGAGCGCTTCTGGGCGCATCTAGTGAAGTGGCGCCCTG 960
 QY 937 CTGCAAGGCGCTGGGCGGCGGTCTCTGATGAGAGCGCTGGAGAGGCTGTGGGCGGAG 996
 DB 961 CTGCAAGGCGCTGGGCGGCGGTCTCTGATGAGAGCGCTGGAGAGGCTGTGGGCGGAG 1020
 QY 997 GCTGTTCCGCTGCTGCTGATGATGAGGCTGATGAGGCTGAGGCGCGCGCGCTG 1056
 DB 1021 GCTGTTCCGCTGCTGCTGATGATGAGGCTGATGAGGCTGAGGCGCGCGCGCTG 1080
 QY 1057 TTGCTGATGAGAGAGAGAGGGGGGCGCGCCGACAGAGGCTCTGATCTAGAGATGAGC 1116
 DB 1081 TTGCTGATGAGAGAGAGAGGGGGGCGCGCCGACAGAGGCTCTGATCTAGAGATGAGC 1140
 QY 1117 AGGATGATCCACCTCCAAAGTCTCCGGGCGACCTTCTCTGGAGAGAGACCATCTCTA 1176
 DB 1141 AGGATGATCCACCTCCAAAGTCTCCGGGCGACCTTCTCTGGAGAGAGACCATCTCTA 1200
 QY 1177 -----C 1177
 DB 1201 CCCCAGAGAGCTGTCACTCTAGCATTTTGAAGAGCTGCGACAGAGACCGGGAAGAGG 1260
 QY 1178 CCGTTGACAGCCCTTCCGACAGAGATGGGCTCTGAGGCTTAACCATTTCCAGCTGAGT 1237
 DB 1261 CCGTTGACAGCCCTTCCGACAGAGATGGGCTCTGAGGCTTAACCATTTCCAGCTGAGT 1320
 QY 1238 TTGCTTCCGAGACCTTCCGACAGAGTGTGCGCCCTTGAAGCTCCGAGAGGGGGGCT 1297
 DB 1321 TTGCTTCCGAGACCTTCCGACAGAGTGTGCGCCCTTGAAGCTCCGAGAGGGGGGCT 1380
 QY 1298 GGGCCTGTATCTCAGAAAGGAGGGGACAGCTACACCTGACCAAAAGCCCTGACAC 1357
 DB 1381 GGGCCTGTATCTCAGAAAGGAGGGGACAGCTACACCTGACCAAAAGCCCTGACAC 1440
 QY 1358 TTGTATCTGATCTTGGGCTGTCTGACAGTGTCAAGGTGACACACTGCTCATGTCA 1417
 DB 1441 TTGTATCTGATCTTGGGCTGTCTGACAGTGTCAAGGTGACACACTGCTCATGTCA 1500
 QY 1418 CACTGCCCCGCTGAGATCTTCCCTGGGCGCTGCGCCCTGAGCTTCCAGACACAT 1477
 DB 1501 CACTGCCCCGCTGAGATCTTCCCTGGGCGCTGCGCCCTGAGCTTCCAGACACAT 1560
 QY 1478 TCTTTGGCCTTAAGGGCTTCTCTCAGAGACCTTAATTTGACCAACAACAACCTGGGCTT 1537
 DB 1561 TCTTTGGCCTTAAGGGCTTCTCTCAGAGACCTTAATTTGACCAACAACAACAACCTGGGCTT 1620
 QY 1538 CAGCCACATCAGTGGGACCTGAGAGCTGGGCTGACATGGGCGCTGCTCACCTTCCCA 1597

DB 1621 CAGCCACATCACTGGGCACTGGAGCTGGGGTGCACATGGGGCTGTCACTTGCCACA 1680
 QY 1598 CATCTTCAGCCAGCCAGGGGCGCTGCCAGCTTCAATTTACAGACCTGACTCTCACCCT 1657
 DB 1681 CATCTTCAGCCAGCCAGGGGCGCTGCCAGCTTCAATTTACAGACCTGACTCTCACCCT 1740
 QY 1658 TCCCCCTGCTGTCCAGAGCTGAACATTAAGCTTGAATGTCACTGAGAGTGTAC 1717
 DB 1741 TCCCCCTGCTGTCCAGAGCTGAACATTAAGCTTGAATGTCACTGAGAGTGTAC 1800
 QY 1718 ATGGAGGTGTATGGAGCATATACAAAGCTACTGTGACATGGGGCCAAACAG 1777
 DB 1801 ATGGAGGTGTATGGAGCATATACAAAGCTACTGTGACATGGGGCCAAACAG 1860
 QY 1778 TAAACAGCCACCTTCTTGAAGAGAGATGCAAAAGCTTTGGGGTGTATGAAAGACCTT 1837
 DB 1861 TAAACAGCCACCTTCTTGAAGAGAGATGCAAAAGCTTTGGGGTGTATGAAAGACCTT 1920
 QY 1838 TTACAATATATACCAATTAACTGCCCTGGAAGAGGCGATAGTGG 1883
 DB 1921 TTACAATATATACCAATTAACTGCCCTGGAAGAGGCGATAGTGG 1966
 RESULT 9
 AAH9646
 ID AAH9646 standard; cDNA; 1570 BP.
 AC XX
 DT 16-OCT-2001 (first entry)
 XX XX
 DE Human protein encoding cDNA sequence SEQ ID NO:481.
 XX XX
 KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antineumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiac; central nervous system; virocid;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;
 KW antiagregant; haemostatic; vulnerary; antitumor; osteopathic; eczema;
 KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antinaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac amyphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW chromocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.
 KW Homo sapiens.
 OS XX
 PN MO200153455-A2.
 XX XX
 PD 26-UTL-2001.
 XX XX
 XX 22-DEC-2000; 2000MO-US035017.
 PF XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PA (HYSB-) HYSBQ INC.
 XX XX
 PI Tang YT, Liu C, Drmanac RT;
 XX XX
 DR WPI; 2001-457603/49.
 DR P-PSDB; AAM25705.
 XX XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
 XX XX
 PS Claim 1; Page 552; 1217bp; English.

XX	Homo sapiens.
FT	Key
FT	Location/Qualifiers
FT	CDS
FT	127..1107
FT	/tag= a
XX	/product= "APRG polypeptide"
PN	MO200192527-A2.
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001W0-US017581.
XX	
PR	01-JUN-2000; 2000US-0209407P.
XX	
PR	30-NOV-2000; 2000US-0250326P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
P1	Tang YT, Azimzai Y, Yue H, Burford N, Ding L, Elliott VS,
P1	Patterson C, Baughn MR;
DR	WPI: 2002-114350/15.
DR	P-PSDB: ABB027263.
PT	Novel human apoptosis regulator polypeptides and polynucleotides for
PT	diagnosing, preventing, treating cell proliferative, immunological and
PT	reproductive disorders and for identifying modulators of therapeutic use
PS	Claim 5; Page 101; 103pp; English.
XX	
CC	The invention provides human apoptosis regulator (APRG) polypeptides and
CC	polynucleotides. The APRG polypeptides, polynucleotides and modulators
CC	are useful for diagnosis, treatment and prevention of cell proliferative
CC	immunological and reproductive disorders. The cell proliferative
CC	disorders include cancers, actinic keratosis, arteriosclerosis,
CC	atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, and
CC	immunological disorders include acquired immunodeficiency syndrome
CC	(AIDS), adult respiratory distress syndrome, Addison's disease,
CC	anaphylaxis, spondylitis, amyloidosis, allergies, anemia, osteoporosis,
CC	autoimmune hemolytic anemia, asthma, autoimmune thyroiditis, Crohn's
CC	disease, contact dermatitis, diabetes mellitus, gout, Graves' disease,
CC	glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
CC	erythematosus, systemic sclerosis, ulcerative colitis, hemodialysis,
CC	uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC	infections and trauma. Reproductive disorders include disorders of
CC	prolactin production, infertility, endometriosis, polycystic ovary
CC	syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,
CC	disruptions of spermatogenesis, cancer of testis and prostate, impotence,
CC	carcinoma of male breast and gynecomastia. The APRG polynucleotides are
CC	useful for creating knock-in humanized animals or transgenic animals to
CC	model human disease and to detect and quantify gene expression in
CC	biopsied tissues in which expression of APRG is correlated with disease.
CC	APRG, fragments of it and antibodies specific for APRG are useful as
CC	elements on a microarray which is useful to monitor or measure protein-
CC	protein interactions, drug-target interactions and gene expression
CC	profiles. The present sequence represents a human APRG polypeptide
CC	encoding CDNA
XX	
SQ	Sequence 1230 BP; 205 A; 382 C; 433 G; 210 T; 0 U; 0 Other;
OY	Query Match 62.7%; Score 1180.8; DB 6; Length 1230;
DB	Best Local Similarity 99.8%; Pred. No. 6,8e-253;
	Matches 1182; Conservative 0; Mismatches 2; Indels 0; Gaps 0
1	AGCGCTAATATATAGAGAAGGTGCCAGAAAGATCAAAAACAGTGGCTGGCGCGTCGCC
4	AGGCGTAATATATATAGAGAAGGTGCCAGAAAGATCAAAAACAGTGGCTGGCGCGTCGCC

OY	61	AGAGATCATATGAGACCGCCAGAAATCTGGGCGGGTTCTGAGCTTTTCCGCGCTCCCTCCCG	120
Db	64	AGGAGTACATGAGACGCGCAAGATCTGGCGGGATTCTGAGCTTTTCCGCGCTCCCTCCCG	123
OY	121	GGAAATGCGCGTATCCGGGTGCAACCCCGCGCCCGTGTGGAGAGAGATGATGTCCTTGAC	180
Db	124	GGAAATGCGCGTATCCGGGTGCAACCCCGCGCCCGTGTGGAGAGAGATGATGTCCTTGAC	183
OY	181	TACTACGGGATGCTGTGTGCTTACCCGTAATGTTGAGGTGGTGGCGGGCAATGACCGAG	240
Db	184	TACTACGGGATGCTGTGTGCTTACCCGTAATGTTGAGGTGGTGGCGGGCAATGACCGAG	243
OY	241	TGCGAGCTGGAGGCTCTGGGCTTTCTGTGGAATGAGGCTCTGGCGCGCGCGAGGCTTA	300
Db	244	TGCGAGCTGGAGGCTCTGGGCTTTCTGTGGAATGAGGCTCTGGCGCGCGCGAGGCTTA	303
OY	301	GCCCGGGCCGCGACGCGCTTAGAGCTTCTGTGAGCTTGAGCGCGCGCGGCACTGCGGC	360
Db	304	GCCCGGGCCGCGACGCGCTTAGAGCTTCTGTGAGCTTGAGCGCGCGCGGCACTGCGGC	363
OY	361	GAGAGCAACTGGCGGCTGTGGGGCAACTCTGCGGCGTGGCTGGCGCGCGCAAGACTGCTG	420
Db	364	GAGAGCAACTGGCGGCTGTGGGGCAACTCTGCGGCGTGGCTGGCGCGCGCAAGACTGCTG	423
OY	421	CCGACCTGGCGCGCAAGCGCGCGCGCACTGTCTCAGAACGCTTAGCTATGACACC	480
Db	424	CCGACCTGGCGCGCAAGCGCGCGCGCACTGTCTCAGAACGCTTAGCTATGACACC	483
OY	481	TCCAGCTTCTTAAGAAGAGACAGAGGGTAAGTCCCGTGGCGGTGACAGCATCTTCT	540
Db	484	TCCAGCTTCTTAAGAAGAGACAGAGGGTAAGTCCCGTGGCGGTGACAGCATCTTCT	543
OY	541	GCAAAATCTCAGCAGGGGTCAGTGGGAGACAGAGCTCCCGCCCAACCAAGCGGACGCGG	600
Db	544	GCAAAATCTCAGCAGGGGTCAGTGGGAGACAGAGCTCCCGCCCAACCAAGCGGACGCGG	603
OY	601	AGTCGGGGCCGGCCAGTGTGTGTGCCAAGCGCGGCTGAGAGGGGCCCCAGCCGCAACC	660
Db	604	AGTCGGGGCCGGCCAGTGTGTGTGCCAAGCGCGGCTGAGAGGGGCCCCAGCCGCAACC	663
OY	661	CAGAGACATGACAGCCCGCCAGACTTCTCTGTAAGGCAAGAGCACTGTGACATCCG	720
Db	664	CAGAGACATGACAGCCCGCCAGACTTCTCTGTAAGGCAAGAGCACTGTGACATCCG	723
OY	721	CTCCGGGTTGAGAGAGATCTGTGAGCATGGGACGCTTGTGAGCAAGGGCGTGGATCC	780
Db	724	CTCCGGGTTGAGAGAGATCTGTGAGCATGGGACGCTTGTGAGCAAGGGCGTGGATCC	783
OY	781	CGCGGGCCCGACGGCGCTGGCGCGCGCACTGAGCGTGTGTTGGGACGGGCACTGCAGTGCTG	840
Db	784	CGCGGGCCCGACGGCGCTGGCGCGCGCACTGAGCGTGTGTTGGGACGGGCACTGCAGTGCTG	843
OY	841	CGCTCAAGGGACCTGGGCTGTGTGTTGTGACATCAAGTTCTAGAGCTCTCTATCTG	900
Db	844	CGCTCAAGGGACCTGGGCTGTGTGTTGTGACATCAAGTTCTAGAGCTCTCTATCTG	903
OY	901	GAGCGCTTCTGGGGCGACTACCTGATGTCGCGCTGTGCAAGGCTTGGCGGGCGTGTTC	960
Db	904	GAGCGCTTCTGGGGCGACTACCTGATGTCGCGCTGTGCAAGGCTTGGCGGGCGTGTTC	963
OY	961	CTGACTAGGCGCTGCGACAGGGCTGTGGGCGCGGAGCGTGTGCGCTGTGGTGAAGTGTG	1020
Db	964	CTGACTAGGCGCTGCGACAGGGCTGTGGGCGCGGAGCGTGTGCGCTGTGGTGAAGTGTG	1023
OY	1021	GATAGAGCTGACTATGAGGCTGGCGCGCGCTGTGTGCTGATGAGAGAGAAAGGGGG	1080
Db	1024	GATAGAGCTGACTATGAGGCTGGCGCGCGCTGTGTGCTGATGAGAGAGAAAGGGGG	1083
OY	1081	CGGCGCCCGACAGAGGCTCTCTGATTCAGAGACTGGCAGAAATTAATCCACTTCCAACTCT	1140
Db	1084	CGGCGCCCGACAGAGGCTCTCTGATTCAGAGACTGGCAGAAATTAATCCACTTCCAACTCT	1143

QY 1141 CC6G6CCACCTTCTCTG6GAGAGACCATCTCTACCCCTTGA 1184
 DB 1144 CC6G6CCACCTTCTCTG6GAGAGACCATCTCTACCCCTTGA 1187

RESULT 11

ADCT9259 standard; cDNA; 981 BP.

AC ADC79259;

DT 01-JAN-2004 (first entry)

DE Human DEDD2 encoding cDNA SEQ ID NO:1.

XX human; death effector domains containing DNA-binding protein;

KM DED2-containing DNA-binding protein; DEDD2; cell death; gene therapy;

KM cytostatic; cancer; chronic myeloid leukemia; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..981

FT /tag= a

FT /product= "DEDD2"

PN MO2003054195-A1.

PD 03-JUL-2003.

XX 20-DEC-2002; 2002MO-JP013371.

XX 20-DEC-2001; 2001JP-00387854.

PR 18-JUL-2002; 2002JP-00209458.

XX (MORG) MORINAGA MILK IND CO LTD.

PA Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;

PI WPI; 2003-5569246/53.

XX P-PSDB; ADCT9260.

DR DNA encoding cell death proteins for treatment of kidney, large intestine

PT and prostate cancers and leukemia.

XX Claim 2; Page 18-20; 26pp; Japanese.

XX The present sequence encodes a human death effector domains (DED)

CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that

CC causes cell death. Also described: (1) primer and probe for investigation

CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has

CC cytostatic activity. DEDD2 can be used in the diagnosis and treatment of

CC cancers of the kidney, large intestine and prostate, and acute and

CC chronic myeloid leukemia.

XX Sequence 981 BP; 151 A; 303 C; 363 G; 164 T; 0 U; 0 Other;

XX Query Match 52.0%; Score 979.4; DB 10; Length 981;

XX Best Local Similarity 99.9%; Pred. No. 4.4e-208;

XX Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 124 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCGAGAGAGATGAGTGGCTGACTAC 183

DB 1 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCGAGAGAGATGAGTGGCTGACTAC 60

QY 184 TACGGAGTGTGCTGCTTCAACCGTATGTTGAGGTGTGTGGCGGCAACTGACGAGTGC 243

DB 61 TACGGAGTGTGCTGCTTCAACCGTATGTTGAGGTGTGTGGCGGCAACTGACGAGTGC 120

QY 244 GAGGTGAGTGTGCTGCTTCAACCGTATGTTGAGGTGTGTGGCGGCAACTGACGAGTGC 303

DB 121 GAGGTGAGTGTGCTGCTTCAACCGTATGTTGAGGTGTGTGGCGGCAACTGACGAGTGC 180

QY 304 CG6G6CCGAGCGG3CTTGAAGCTTCTG6GAGCTGAGAGCGCGCGG6CAGTCCGCGAG 363
 DB 181 CG6G6CCGAGCGG3CTTGAAGCTTCTG6GAGCTGAGAGCGCGCGG6CAGTCCGCGAG 240

QY 364 AGCAACTGCGGCTGCTGCGG6CAACTCTGCGG6CAGTCCGCGG6CAGTCCGCGAG 423

DB 241 AGCAACTGCGGCTGCTGCGG6CAACTCTGCGG6CAGTCCGCGG6CAGTCCGCGAG 300

QY 424 CACTGCGGCGAGCGGCGCGG6CAGTCTTCAAGAGCTTATGACTATGCACTTCC 483

DB 301 CACTGCGGCGAGCGGCGCGG6CAGTCTTCAAGAGCTTATGACTATGCACTTCC 360

QY 484 AGCTTTCAAG 543

DB 361 AGCTTTCAAG 420

QY 544 AATTCTCAG 603

DB 421 AATTCTCAG 480

QY 604 CG6G6CCGAGCGG3CTTGAAGCTTCTG6GAGCTGAGAGCGCGCGG6CAGTCCGCGAG 663

DB 481 CG6G6CCGAGCGG3CTTGAAGCTTCTG6GAGCTGAGAGCGCGCGG6CAGTCCGCGAG 540

QY 664 CAGCAGTCAAGAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723

DB 541 CAGCAGTCAAGAGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

QY 724 CGG6GTTGAG 783

DB 601 CGG6GTTGAG 660

QY 784 CGG6GTTGAG 843

DB 661 CGG6GTTGAG 720

QY 844 TCAAGGAGCTGAG 903

DB 721 TCAAGGAGCTGAG 780

QY 904 GCTTCTGAG 963

DB 781 GCTTCTGAG 840

QY 964 ACTGAGGCTGAG 1023

DB 841 ACTGAGGCTGAG 900

QY 1024 GAGGCTGACTATGAG 1083

DB 901 GAGGCTGACTATGAG 960

QY 1084 CGCCCGACAG 1104

DB 961 CGCCCGACAG 981

RESULT 12

ABL39693 standard; cDNA; 1106 BP.

AC ABL39693;

DT 10-MAY-2002 (first entry)

DE Human NS cDNA sequence SEQ ID NO:3.

XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;

KM antineoplastic; antiarthritic; antipruritic; ophthalmological; anti-HIV;

KM vasorelaxant; antidiabetic; antidiabetic; antidiabetic; antidiabetic;

KM anorectic; muscular; antidiabetic; antidiabetic; antidiabetic; antidiabetic;

KM antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;

KM antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;

KM gastrointestinal; virucide; antiulcer; cerebroprotective; nootropic;
 KM contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KM endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KM rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KM inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KM infertility; cardiovascular disease; coagulation disease; hypertension;
 KM ischemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KM diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KM gastric ulcer; Alzheimer's disease; gene; ss.
 OS Homo sapiens.
 XX MO200206315-A2.
 XX 24-JAN-2002.
 PD 17-JUL-2001; 2001MO-IL000653.
 PF 18-JUL-2000; 2000IL-00137345.
 XX 15-DEC-2000; 2000IL-00140354.
 PR (COMP-) COMPUGEN LTD.
 XX
 PA Minz L, Freilich S, Bernstein J;
 PI WPI; 2002-155037/20.
 XX P-PSDB; ABB06039.
 DR
 XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 PT
 XX Claim 1, Page 77; 290pp; English.
 XX
 XX ABLJ3691 to ABLJ3818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antidiabetic, antipsoriatic, ophthalmological, virucide,
 CC vasorelaxant, antihypertensive, antidiabetic, dermatological,
 CC vasorelaxant, muscular, anti-HIV, antileukemic, cardiovascular,
 CC antitumor, antifibrotic, hypotensive, antidiabetic, tranquilizer, anti-
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilizer, anti-
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC nootropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dyslexia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive
 XX
 SQ Sequence 1106 BP; 182 A; 340 C; 385 G; 189 T; 0 U; 10 Other;
 Query Match 50.3%; Score 946.4; DB 6; Length 1106;
 Best Local Similarity. 97.7%; Pred. No. 1e-200;
 Matches 1036; Conservative 7; Mismatches 8; Indels 9; Gaps 8;

Db 189 TATGATCGAGTGTGTGGGCGCACTGACCGAGTCRAGCTGAGCTTCTGCTTTCT 248
 QY 267 GCTGATGAGGCTCTCTGGCGCCGCGAGGCTTAGCCGGGCGCCGAGCGCTTAGAGCT 326
 Db 249 GCTGATGAGGCTCTCTGGCGCCGCGAGGCTTAGCCGGGCGCCGAGCGCTTAGAGCT 307
 QY 327 CCTGCTGAGCTGAGAGCGCGCGGAGTGTGGCGAGCAACTGCGGCTCTGGGCGCA 386
 Db 308 CCTGCTGAGCTGAGAGCGCGCGGAGTGTGGCGAGCAACTGCGGCTCTGGGCGCA 367
 QY 387 ACTCCGCGCGTGTGGCGCGCGGAGTGTGGCGAGCAACTGCGGCGCGGCGCGCGCG 446
 Db 368 ACTCCGCGCGTGTGGCGCGCGGAGTGTGGCGAGCAACTGCGGCGCGGCGCGCGCG 425
 QY 447 GCCAGTGTCTCCAGAGCGCTATAGCTATGAGCACTTCAAGCTTTCAAGAGAGAGAGG 506
 Db 426 GCCAGTGTCTCCAGAGCGCTATAGCTATGAGCACTTCAAGCTTTCAAGAGAGAGAGG 485
 QY 507 TAGCTGCGGTCGCGCGGAGTCAAGAGTGTCTGCAATTTCTCAGCAGGCTCAGTGGGA 566
 Db 486 TAGCTGCGGTCGCGCGGAGTCAAGAGTGTCTGCAATTTCTCAGCAGGCTCAGTGGGA 545
 QY 567 GACAGGCTCCCGCCCAACCAAGCGGAGCGGAGTGTGGCGCGGCGCGCGCGCGGCG 626
 Db 546 GACAGGCTCCCGCCCAACCAAGCGGAGCGGAGTGTGGCGCGGCGCGCGCGCGGCG 604
 QY 627 CAGACGCGCGGAGAGGCGGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 686
 Db 605 CAGACGCGCGGAGAGGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 663
 QY 687 TTCTCTGAGGCAAGTGAAGTGTGACATCCGCGCTCGGGTTTCAAGCAGAGTACTGCA 746
 Db 664 TTCTCTGAGGCAAGTGAAGTGTGACATCCGCGCTCGGGTTTCAAGCAGAGTACTGCA 723
 QY 747 GATGAGCGGAGGCTTGAAGAGGCGTGTGACATCCGCGCGCGCGCGCGCGCGCGGCA 806
 Db 724 GATGAGCGGAGGCTTGAAGAGGCGTGTGACATCCGCGCGCGCGCGCGCGCGGCA 782
 QY 807 GCTGAGCGGTTTGGGCGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 866
 Db 783 GCTGAGCGGTTTGGGCGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGG 842
 QY 867 TTGTGACATCAAGTGTGACAGCTCTCTATATGAGCGCTTCTGAGGCGGAGTACTGAG 926
 Db 843 TTGTGACATCAAGTGTGACAGCTCTCTATATGAGCGCTTCTGAGGCGGAGTACTGAG 902
 QY 927 TGGCGCGCGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 986
 Db 903 TGGCGCGCGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 962
 QY 987 GGGCGGAGGCGGTTTGGCGGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1046
 Db 963 GGGCGGAGGCGGTTTGGCGGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1020
 QY 1047 GGGCGGCGGTTTGGCGGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1086
 Db 1021 GGGCGGCGGTTTGGCGGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1060

RESULT 13
 AAF27407
 ID AAF27407 standard; cDNA; 909 BP.
 XX
 AC AAF27407;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Human apoptosis-associated factor NT2RM1000558 partial cDNA, SEQ ID NO.1.
 XX
 KW Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
 KW DBD; caspase family cleavage domain; pro-apoptotic; drug screening;
 KW cell proliferation; ischemic disease; chronic viral disease; ss.

XX Homo sapiens.
 OS MO200104300-A1.
 XX 18-JAN-2001.
 XX 06-JUL-2000; 2000MO-JP004516.
 XX 08-JUL-1999; 99JP-00194179.
 XX 18-OCT-1999; 99US-0159586P.
 XX (HELI-) HELIX RES INST.
 XX Oca T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Sato S;
 XX WPI, 2001-118348/14.
 XX P-PSDB; AAB60386.
 XX Polynucleotide encoding an apoptosis-associated factor protein with death
 PT effector domain and caspase family-cleavage domain, useful in regulating
 PT diseases with cell proliferation.
 XX Claim 1; Page 41-43; 53pp; Japanese.
 XX The invention relates to a novel human apoptosis-associated factor
 CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
 CC effector domain (DED) and a caspase family cleavage domain and is capable
 CC of inducing apoptosis in cells. The invention also relates to nucleic
 CC acids encoding the protein (AA27407, AA27408); variants of the protein
 CC (particularly dominant negative variants); vectors and host cells
 CC comprising a nucleic acid which encodes an apoptosis-associated factor
 CC of the invention; the recombinant production of the protein; an antibody
 CC against the protein; and methods of screening for compounds which can
 CC regulate apoptosis. The apoptosis-related factor is useful in regulating
 CC diseases associated with cell proliferation and in screening drug
 CC candidates e.g., for regulating cell proliferation or cell death in
 CC ischaemic diseases and chronic viral diseases. The present sequence
 CC represents cDNA encoding a substantial proportion of the human apoptosis-
 CC associated factor NT2RM1000558
 XX
 XX Sequence 909 BP; 139 A; 284 C; 332 G; 154 T; 0 U; 0 Other;
 SQ
 Query Match 48.3%; Score 909; DB 5; Length 909;
 Best Local Similarity 100.0%; Pred. No. 2e-192;
 Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 124 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCTGGAGAGATGAGTCTGACTAC 183
 DB 1 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCTGGAGAGATGAGTCTGACTAC 60
 QY 184 TAGGGATGTCCTGCTTCAACCGTATGAGTGTGGCGGGCAATGACCGAGTGC 243
 DB 61 TAGGGATGTCCTGCTTCAACCGTATGAGTGTGGCGGGCAATGACCGAGTGC 120
 QY 244 GACCTGAGAGCTCTGAGCTTCTGCTGATGAGGCTCTGAGCGCGCGGAGCTTACC 303
 DB 121 GACCTGAGAGCTCTGAGCTTCTGCTGATGAGGCTCTGAGCGCGCGGAGCTTACC 180
 QY 304 CGGGCCCGGACCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGCTTACC 363
 DB 181 CGGGCCCGGACCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGCTTACC 240
 QY 364 AGGAACCTGGGGCTGCTGGGGCAATCTGCTGCTGGCGCGCGGAGCTTACC 423
 DB 241 AGGAACCTGGGGCTGCTGGGGCAATCTGCTGCTGGCGCGCGGAGCTTACC 300
 QY 424 CACCTGGCGGACCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGGAGCTTACC 483
 DB 301 CACCTGGCGGACCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGGAGCTTACC 360
 QY 484 AGCTTTCAAGAGAGAGAGGAGTGTGCTGCTGCTGGCGCGGAGCTTACC 543

DB 361 AGCTTTCAAGAGAGAGAGGAGTGTGCTGCTGCTGGCGCGGAGCTTACC 420
 QY 544 AATTCTCAGAGAGGTCAGTGGAGAGAGGCTTCCCGCCCAACCAAGCGGAGCGGAGT 603
 DB 421 AATTCTCAGAGAGGTCAGTGGAGAGAGGCTTCCCGCCCAACCAAGCGGAGCGGAGT 480
 QY 604 CGGGCCCGGACCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGCTTACC 663
 DB 481 CGGGCCCGGACCGGCTTGAAGCTTCTGCTGAGAGTGGAGCGCGCGGAGCTTACC 540
 QY 664 CAGCAGTCAAGAGCGCGGACCGGCTTCTGCTGAGAGTGGAGCGCGGAGCTTACC 723
 DB 541 CAGCAGTCAAGAGCGCGGACCGGCTTCTGCTGAGAGTGGAGCGCGGAGCTTACC 600
 QY 724 CGGGTTCAAGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 783
 DB 601 CGGGTTCAAGAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 660
 QY 784 CGGCCCCAGGCGCTGAGCGCGGAGCTGAGAGTGTGGAGCGGAGCTTACC 843
 DB 661 CGGCCCCAGGCGCTGAGCGCGGAGCTGAGAGTGTGGAGCGGAGCTTACC 720
 QY 844 TCAAGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
 DB 721 TCAAGGAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 QY 904 GCCTTCGAGGAGGACATCACTGAGTGGAGCGGCTGCTGAGAGCGGAGCGGAGTCTG 963
 DB 781 GCCTTCGAGGAGGACATCACTGAGTGGAGCGGCTGCTGAGAGCGGAGCGGAGTCTG 840
 QY 964 ACTGAGGCGCTGCGAGAGGCTGTGGCGGGAGGCTGTGCTGCTGCTGCTGCTGCTG 1023
 DB 841 ACTGAGGCGCTGCGAGAGGCTGTGGCGGGAGGCTGTGCTGCTGCTGCTGCTGCTG 900
 QY 1024 GAGGCTGAC 1032
 DB 901 GAGGCTGAC 909
 RESULT 14
 ABL39692
 ID ABL39692 standard; cDNA; 1084 BP.
 XX
 AC ABL39692;
 XX
 DT 10-MAY-2002 (first entry)
 XX
 DE Human NS cDNA sequence SEQ ID NO:2.
 XX
 KW Human; cytosolic; osteopathic; gynaecological; neuroprotective;
 KW antihemetic; antihemetic; antihemetic; antihemetic; anti-HIV;
 KW vasotropic; antihemetic; antihemetic; antihemetic; dermatological;
 KW anorectic; muscular; antihemetic; cardiovascular; antihemetic;
 KW antihemetic; antihemetic; antihemetic; antihemetic; antihemetic;
 KW gastroenteric; antihemetic; antihemetic; antihemetic; antihemetic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dyslexia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; retinosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW infertility; cardiovascular disease; coagulation disease; hypertension;
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease; gene; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO200206315-A2.
 XX
 XX 24-JAN-2002.
 XX
 XX 17-JUL-2001; 2001MO-11000653.
 XX

PR 18-JUL-2000; 2000IL-00137345.
 PR 15-DEC-2000; 2000IL-00140354.
 PA (COMP-) COMFUGEN LTD.
 XX
 XX
 PI Mintz L, Freilich S, Bernstein J;
 DR WPI, 2002-155037/20.
 DR P-PDB; ABB06038.
 XX
 XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 PT
 PS Claim 1; Page 76-77; 290pp; English.

CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virocidic,
 CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antiinfectivity, cardiovascular,
 CC anticoagulant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,
 CC immunorepressant, gastrointestinal, antileptotic, cerebroprotective,
 CC neurotropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dyslexia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, retinosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC depression, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive

XX Sequence 1084 BP; 174 A; 332 C; 378 G; 190 T; 0 U; 10 Other;

Query Match 47.9%; Score 902.4; DB 6; Length 1084;
 Best Local Similarity 97.2%; Pred. No. 6.2e-191;
 Matches 984; Conservative 7; Mismatches 13; Indels 8; Gaps 7;

QY 75 GGCAGATCTGGCGCGGTTCTGAGCTTGTCCCTCCCTCCCGGGAATGGCGCTATC 134
 DB 35 GGCAGACCTTGTGGCGGTTCTGAGCTTGTCCCTCCCTCCCGGGAATGGCGCTATC 94
 QY 135 CGGCTCGACCCCGCGCGCTGCTGGAGAGAGATGAGTGGCTGGACTAAGCGAGTGC 194
 DB 95 CGGCTCGACCCCGCGCGCTGCTGGAGAGAGATGAGTGGCTGGACTAAGCGAGTGC 154
 QY 195 GTGCTTACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGCAGCTGAGCT 254
 DB 155 GTGCTTACCGTATGTTGAGGTGTGGCGGGCAACTGACCGAGTGCAGCTGAGCT 214
 QY 255 CCGTGGCTTTCTGCTGATGATAGGCTCTGAGCGCGCGGAGGCTTGAAGCGCGGAG 314
 DB 215 CCGTGGCTTTCTGCTGATGATAGGCTCTGAGCGCGCGGAGGCTTGAAGCGCGGAG 273
 QY 315 CGGCTGAGAGCTCTGCTGAGAGTGGAGCGCGCGGAGTGGCGGAGAGCAACTGGCG 374
 DB 274 CGGCTGAGAGCTCTGCTGAGAGTGGAGCGCGCGGAGTGGCGGAGAGCAACTGGCG 333
 QY 375 GCTGCTGGGCAACTCTGCGCGTGTGCGCGCCGCAAGACTTGTGCGGCACTTGGCGG 434
 DB 334 GCTGCTGGGCAACTCTGCGCGTGTGCGCGCCGCAAGACTTGTGCGGCACTTGGCGG 393
 QY 435 CAAGCGCGCGCGCGCGAGTGTCTCCAGAACCTTATAGTATGGCACTCCAGCTTTCAAA 494
 DB 394 CAA--AGCGCGCGCGCGAGTGTCTCCAGAACCTTATAGTATGGCACTCCAGCTTTCAAA 451
 QY 495 GAGGAGAGAGGAGTGGCGCGTGGCGGTGGGAGTCAAGAGATTGCAAAATTCAGCA 554
 DB 452 GAGGAGAGAGGAGTGGCGCGTGGCGGTGGGAGTCAAGAGATTGCAAAATTCAGCA 511

QY 555 GGGTACGTGGAGACAGGCTCTCCCGCCCAACAGCGGCGCGCGGAGTGGCGCGCGGC 614
 DB 512 GGGTACGTGGAGAGAGGCTCTCCCGCCCAACAGCGGCGCGCA-CGCGAGTGGCGCGCGGC 570
 QY 615 CAGTGGTGTGTCACAGCGCGCGCGGAGAGGGGGCCCGAGCGCGGACCCAGCAGCAGTCA 674
 DB 571 CAGTGGTGTGTCACAGCGCGCGCGGAGAG-GGGCCCGAGCGCGGACCCAGCAGCAGTCA 629
 QY 675 GCCCGCAGACCTTCTCTGAGAGGCAAGTGAAGTGAATCCAGTCCGAGTTTCAGC 734
 DB 630 GCCCGCAGACCTTCTCTGAGAGGCAAGTGAAGTGAATCCGAGTTTCAGC 689
 QY 735 AAGATCTGCGAGCATGGGCCAGCTTGGAGAGAGGCGGTGGATCCCGCGCGCCAGGC 794
 DB 690 AAGATCTGCGAGCATGGGCCAGCTTGGAGAGAGGCGGTGGATCCCGCGCGCCAGGC 748
 QY 795 GCGGCGCGCGCGAGCTGGAAGTGTGGGAGGCGCACCGGAGTGGCGTCAAGGAGACT 854
 DB 749 GCTGGGCGCGCGCAGGAGTGTGGGAGGCGCACCGGAGTGGCGTCAAGGAGACT 808
 QY 855 GGGCTCTGTGTTGTGACATCAAGTTCAGAGCTTCTTATCTGACAGCTTGTGGG 914
 DB 809 GGGCTCTGTGTTGTGACATCAAGTTCAGAGCTTCTTATCTGACAGCTTGTGGG 868
 QY 915 CCACTACTGAGTGGCGCTTGTCTGAGGCGCTTGGCGGCGGTGTTCTGACTGAGGCGCT 974
 DB 869 CCACTACTGAGTGGCGCTTGTCTGAGGCGCTTGGCGGCGGTGTTCTGACTGAGGCGCT 928
 QY 975 GGGAGAGGCTGGGCGCGGAGAGGCTGTTGCGCTGCGGTGAGTGGAGGCTGACTA 1034
 DB 929 GGGAGAGGCTGGGCGCGGAGAGGCTGTTGCGCTGCGGTGAGTGGAGGCTGACTA 987
 QY 1035 TGAGGCTGGCGCGCGCGCTTGTCTGATGAGAGAGAGGCGGCGCGC 1086
 DB 988 TGAGGCTGGCGCGCGCGCTTGTCTGATGAGAGAGAGGCGGCGCGCAC 1038

RESULT 15
 ADI30856
 ID ADI30856 standard; cDNA; 847 BP.
 XX
 AC ADI30856;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Human cDNA #182.
 XX
 KW Human; gene; ss; immunological response; immunopathological condition;
 KW Crohn's disease; asthma; ulcerative colitis; hyperosinophilia;
 KW irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
 KW acute monocytic leukemia; antiinflammatory; antiaesthetic; antiulcer;
 KW osteopathic; antiarthritic; antirheumatic; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN US6607879-B1.
 XX
 PD 19-AUG-2003.
 XX
 PF 09-FEB-1998; 98US-00023655.
 XX
 PR 09-FEB-1998; 98US-00023655.
 XX
 PA (INCY-) INCYTE CORP.
 XX
 PI Cocks BG, Stuart SG, Seilhamer JI;
 PT WPI, 2003-895307/82.
 PT A composition comprising a plurality of cDNAs, useful for detecting
 PT altered expression of genes in an immunological response to a
 PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma

PT or osteoarthritis.

PS Claim 1; SEQ ID NO 182; 50pp; English.

CC The invention relates to a composition comprising a plurality of cDNAs
CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypersplenophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 847 BP; 179 A; 277 C; 207 G; 184 T; 0 U; 0 Other;

Query Match	38.2%	Score 719;	DB 11;	Length 847;
PostgreSQL 9.5.1	0.0%	Score 0	DB 0	Length 0

Matches 788; Conservative 0; Mismatches 0;

Matches 788; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

1096 GCCTCCTGATCCAGGACTGGCAGGATTGATCCCACTCCAAGTCTCCGGGCCACTTCTC 1155

Db 1 GCCTCCTGATCCAGGACTGGCAGGATTGATCCCACTCCAAGTCTCCGGGCCACCTTCTC 60

QY 1156 CTGGAGGACGACCATCTCTA----- 1176

DB 61 CTGGAGGACGACCACTCTACCCCTAGAGGACTGTCACTCTAGCACTTTGAGGACTGC 120

11 / -----CCCCCTTGACAGCCCTCCACAGGATGTGGCTTGAGGC 1216

DD 121 GACAGGACCGGACAGCAGCCLL1GACAGCCLL1CCACAGGAG1GGCL1GAGGC 180

xy	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
xy	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 104

[illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

[illegible]

1457 1516

421 480

1517 GACCACAACCAACTTGGGCTTCAGCCACATCAGTGGGCACTGGAGCTGGGTGCACATGG 1576

Db 481 GACCACACCACTGGCCTTACGCCATCAGTGGCACTGGAGCTGGGCTGCATGG 540

1577 GGCCTGCTCACCTTGCCCAACATCTCCAGCCAGCGCCCTGCCAGCTTCAATTTA 1636

D5 541 GGCCTGCTCACCTTGCCCAACACATCTCCAGCCAGCCAGGCCCTGCCCAGCTTCAATTTA 600

1637 CAGACCTGACTCTCCCTCACCTTCCCCCTGGTGTCCAGAGCTGAACATAGACTTGCACCTT 1696

DB 601 CAGACCTGACCTCTCCCTACCTTCCCCCTGCTGTCCAGAGCTGAACTATGACCTTGCACCTT 660

1097 00A1G1CACC1GAG1G1CA1GAG1G1A1GAG1CA1AC1CAAG1AC1G1 1/30

[illegible][illegible]

.....

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

Search completed: March 23, 2005, 09:53:08
Job time : 1103.55 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 09:03:21 ; Search time 6747.64 Seconds
(without alignments)
10622.233 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 1883
Sequence: 1 aggcgcataatagagaag.....ctggaaggcagagtg995 1883

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapept 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_hlc: *
4: gb_esc3: *
5: gb_esc4: *
6: gb_esc5: *
7: gb_esc6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1736.4	92.2	1866	3	CR619301 full-leng
2	1730.4	91.9	1860	3	CR593642 full-leng
3	1717.4	91.2	1894	3	CR625070 full-leng
4	1354.4	71.9	1510	3	CR599685 full-leng
5	1348.8	71.6	1586	3	CR590827 full-leng
6	1051.4	55.8	1053	3	CR604307 full-leng
7	952.4	50.6	1046	5	CR604307 full-leng
8	896.4	47.6	1021	5	CR604307 full-leng
9	853.4	45.3	1091	5	CR604307 full-leng
10	848.6	45.1	951	5	CR604307 full-leng
11	836.4	44.4	1166	5	CR604307 full-leng
12	829.2	44.0	1027	5	CR604307 full-leng
13	805.2	42.8	1090	5	CR604307 full-leng
14	782.8	41.6	1089	5	CR604307 full-leng
15	775	41.2	993	5	CR604307 full-leng
16	774.8	41.1	837	4	CR604307 full-leng
17	768	40.8	1008	5	CR604307 full-leng
18	764	40.6	1023	5	CR604307 full-leng
19	764	40.6	1023	5	CR604307 full-leng
20	760	40.4	1051	5	CR604307 full-leng
21	759.4	40.3	872	4	CR604307 full-leng
22	750.2	39.8	1161	5	CR604307 full-leng
23	748.2	39.7	931	5	CR604307 full-leng
24	747.8	39.7	1061	5	CR604307 full-leng

25	732.4	38.9	1062	5	BM908148
26	729.8	38.8	1042	5	BQ072022
27	728	38.7	926	4	BG757042
28	719.2	38.2	874	1	AL519410
29	718.2	38.1	993	4	BM559141
30	713.6	37.9	904	5	BM528274
31	712.2	37.8	718	4	BM117988
32	711.2	37.8	820	5	BM158001
33	709.8	37.7	878	4	BM756130
34	708.2	37.6	891	5	BM458265
35	703.4	37.4	732	6	CB321808
36	700.2	37.2	905	4	BM1910528
37	697.4	37.0	4598	3	AK043908
38	697	37.0	1641	3	AK010701
39	696.8	37.0	935	5	BM458264
40	693	36.8	1032	1	AL527221
41	690.8	36.7	947	4	BM1910416
42	683.6	36.3	918	5	BM191631
43	681.4	36.2	924	5	BM442929
44	673	35.7	981	4	BM008603
45	671.2	35.6	783	5	BM602002

ALIGNMENTS

RESULT 1
LOCUS CR619301
DEFINITION full-length cDNA clone CSDDC015YJ24 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION CR619301
VERSION CR619301.1 GI:50500108
KEYWORDS HTC, CNSLT_CDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1866)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1866)
REFERENCE 2 (bases 1 to 1866)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqret@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
Location/Qualifiers
1..1866
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDC015YJ24"
/issue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 92.2%; Score 1736.4; DB 3; Length 1866;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1866; Conservative 0; Mismatches 1; Indels 59; Gaps 1;
QY 49 CGGCGGTGCGCCAGAGTATCGAGCGCGGAGTCTGAGCTTGTTCGC 108
Db 1 CGGCGGTGCGCCAGAGTATCGAGCGCGGAGTCTGAGCTTGTTCGC 60

```

QY 109 CTCCTCTCCCGGGAATGGCGCTATCGGCTGACCCCGGCCCCGTCGGAGAGAT 168
DB 61 CTCCTCTCCCGGGAATGGCGCTATCGGCTGACCCCGGCCCCGTCGGAGAGAT 120
QY 169 GAGTGCCTGACACTACAGGAGATGCTGCTTCAACGATATGTTGCAAGATGAGGCGG 228
DB 121 GAGTGCCTGACACTACAGGAGATGCTGCTTCAACGATATGTTGCAAGATGAGGCGG 180
QY 229 CAATCAGCAGATGAGCTGAGCTCTGAGCTTCTGCTGATAGAGCTCTGAGCC 288
DB 181 CAATCAGCAGATGAGCTGAGCTCTGAGCTTCTGCTGATAGAGCTCTGAGCC 240
QY 289 GCGGAGGCTTAAAGCCCGGCGCGAGCTGAGCTCTGCTGATAGAGCTCTGAGCC 348
DB 241 GCGGAGGCTTAAAGCCCGGCGCGAGCTGAGCTCTGAGCTCTGAGCTCTGAGCC 300
QY 349 GGGCAATGGCGGAGAGCAACCTGCGGCTGAGGCAACTCTGCGCGCTGAGCC 408
DB 301 GGGCAATGGCGGAGAGCAACCTGCGGCTGAGGCAACTCTGCGCGCTGAGCC 360
QY 409 CAGCACTGCTGCGGACCTGCGCGCAAGCGCGCGAGCTCTCAGAACCTAT 468
DB 361 CAGCACTGCTGCGGACCTGCGCGCAAGCGCGCGAGCTCTCAGAACCTAT 420
QY 469 AGCTATGGAACCTCAGCTCTTCAAAAGAGACAGAGGTAAGCTGCGCTGCGCAG 528
DB 421 AGCTATGGAACCTCAGCTCTTCAAAAGAGACAGAGGTAAGCTGCGCTGCGCAG 480
QY 529 TCAAGAGATTTCGAATTTCTCAGAGGCTCAGTGGAGAGCAAGGCTCCCCCAACAG 588
DB 481 TCAAGAGATTTCGAATTTCTCAGAGGCTCAGTGGAGAGCAAGGCTCCCCCAACAG 540
QY 589 CGGCAAGCGGAGAGTGGGCGCGGCGCAGTGTGTGTGCGAGAGCGGCGAGAGGCGC 648
DB 541 CGGCAAGCGGAGAGTGGGCGCGGCGCAGTGTGTGTGCGAGAGCGGCGAGAGGCGC 600
QY 649 CGAGCGGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
DB 601 CGAGCGGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 709 TGGAGATCCGGGCTCGGGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
DB 661 TGGAGATCCGGGCTCGGGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 769 GGGGTGAGATCCCGGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
DB 721 GGGGTGAGATCCCGGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 829 ACCGAGAGTGGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 888
DB 781 ACCGAGAGTGGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 889 CTCTCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 948
DB 841 CTCTCTATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 949 CGGAGCGGTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
DB 901 CGGAGCGGTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 1009 CTGATCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
DB 961 CTGATCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1069 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
DB 1021 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1129 ACCTCCAAAGTCTCGGCGCACTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1176
DB 1081 ACCTCCAAAGTCTCGGCGCACTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 1140

```

```

QY 1177 -----CCCTTGAAGCC 1189
DB 1141 TGTCACTTACGATCTTTGAGAGACTGCGACAGAGACCGGACAGAGCCCTTGAAGCC 1200
QY 1190 CCTCCACAGAGATGAGGCTTGAAGGCTTAACATTTTCAGAGTATGTTCCCTCCAGA 1249
DB 1201 CCTCCACAGAGATGAGGCTTGAAGGCTTAACATTTTCAGAGTATGTTCCCTCCAGA 1260
QY 1250 CTCCTCTAACCCCAAGTGTGCCCCCTTAAAGCTCCGGAAGCGGAGGCTGAGCTATCT 1309
DB 1261 CTCCTCTAACCCCAAGTGTGCCCCCTTAAAGCTCCGGAAGCGGAGGCTGAGCTATCT 1320
QY 1310 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
DB 1321 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1370 TCTTGGGCTGTGTGACAGTGTCAAGAGTGCACACTGCTCATGCTCAGTCCCTG 1429
DB 1381 TCTTGGGCTGTGTGACAGTGTCAAGAGTGCACACTGCTCATGCTCAGTCCCTG 1440
QY 1430 TGAATCTTCCCTGAGGCTTGTGCCCCCTGAGCTGTGCTTCCAGACACACTTCTTGGCTTAA 1489
DB 1441 TGAATCTTCCCTGAGGCTTGTGCCCCCTGAGCTGTGCTTCCAGACACACTTCTTGGCTTAA 1500
QY 1490 GGGCTTCTCTCAGAGAGCTTAAATTTGACCAACCAACCTTGGGCTTGAAGCACTAG 1549
DB 1501 GGGCTTCTCTCAGAGAGCTTAAATTTGACCAACCAACCTTGGGCTTGAAGCACTAG 1560
QY 1550 TGGGCACTGAGAGTGGGAGTGCACATGAGGAGCTGCTCAGCTTCCAGACATCTCAGCCA 1609
DB 1561 TGGGCACTGAGAGTGGGAGTGCACATGAGGAGCTGCTCAGCTTCCAGACATCTCAGCCA 1620
QY 1610 GCCAGGAGCCCTGCGCCAGCTTCAATTTACAGAGCTGACTCTCTCACCTTCCCTGCTG 1669
DB 1621 GCCAGGAGCCCTGCGCCAGCTTCAATTTACAGAGCTGACTCTCTCACCTTCCCTGCTG 1680
QY 1670 TCCAGAGCTGAACATAGATCTTGCACCTTGAATGACCTGAGAGTGCACATAGGAGTGA 1729
DB 1681 TCCAGAGCTGAACATAGATCTTGCACCTTGAATGACCTGAGAGTGCACATAGGAGTGA 1740
QY 1730 TGGCAGCATCATACCAAGGCTTACTTGTGACATGAGGAGCAAAACAGTAAACAGCAC 1789
DB 1741 TGGCAGCATCATACCAAGGCTTACTTGTGACATGAGGAGCAAAACAGTAAACAGCAC 1800
QY 1790 TTCTTGAAGAGAGATGCAAGGCTTGTGGGCTGATGAAAAGACCTTTTACAAATGATA 1849
DB 1801 TTCTTGAAGAGAGATGCAAGGCTTGTGGGCTGATGAAAAGACCTTTTACAAATGATA 1860
QY 1850 CCAATT 1855
DB 1861 CCAATT 1866

```

RESULT 2
 CR593642 1860 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0DK010YK16 of HeLa cells Cot 25-normalized
 DEFINITION of Homo sapiens (human).
 CR593642
 VERSION CR593642.1 GI:50474449
 ACCESSION
 KEYWORDS HTC; CNSLT_CDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1860)
 AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polyes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 CONTACT : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue
 REFERENCE 2 (bases 1 to 1860)

AUTHORS Genome.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genome - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK010YK16"
/tissue="Hela cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 91.9%; Score 1730.4; DB 3; Length 1860;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 1; Indels 59; Gaps 1;
QY 54 GTGCGCCAGAGATGATCGAGCGCAGAAATCTGCGCGGTTCTGAGCTTGTCCGCTTCC 113
Db 1 GTGCGCCAGAGATGATCGAGCGCAGAAATCTGCGCGGTTCTGAGCTTGTCCGCTTCC 60
QY 114 TCCCGCGGGAATGGCGCTATCCGGGTGACCCCGGCGCGCTGCTGGAGAGAGATGAGTG 173
Db 61 TCCCGCGGGAATGGCGCTATCCGGGTGACCCCGGCGCGCTGCTGGAGAGAGATGAGTG 120
QY 174 CTTGAGATCTAGTACGGAGTGTGTGCTTACCGTATGTTGAGGAGTGTGGCGGAGCACT 233
Db 121 CTTGAGATCTAGTACGGAGTGTGTGCTTACCGTATGTTGAGGAGTGTGGCGGAGCACT 180
QY 234 GACCGAGTGCAGAGTGTGAGCTCTTGACCTTCTGAGATGAGGCTCTTGCGCGCGCG 293
Db 181 GACCGAGTGCAGAGTGTGAGCTCTTGACCTTCTGAGATGAGGCTCTTGCGCGCGCG 240
QY 294 AGGCTTAGCCCCGGGCGCGCAGCGGCTTAGAGCTCTGCTGAGAGCTGAGCGCGCGGCA 353
Db 241 AGGCTTAGCCCCGGGCGCGCAGCGGCTTAGAGCTCTGCTGAGAGCTGAGCGCGCGGCA 300
QY 354 GTGCGGCGAGAGCAACTGCGGCTGTGCGGCAACTCTGCGGCTGTGCGGCGCGCA 413
Db 301 GTGCGGCGAGAGCAACTGCGGCTGTGCGGCAACTCTGCGGCTGTGCGGCGCGCA 360
QY 414 CTTGCTGCGCACTGTGCGGCGCAAGCGCGCGCGAGTCTTCCAGAACGCTATAGCTA 473
Db 361 CTTGCTGCGCACTGTGCGGCGCAAGCGCGCGCGAGTCTTCCAGAACGCTATAGCTA 420
QY 474 TGGCACTTCCAGCTCTTCAAGAGAGACAGAGGTAAGTGTGCGGCTGTGCGGCGAG 533
Db 421 TGGCACTTCCAGCTCTTCAAGAGAGACAGAGGTAAGTGTGCGGCTGTGCGGCGAG 480
QY 534 CAGTTTGCAGAAATTTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCCAACCAAGGGCA 593
Db 481 CAGTTTGCAGAAATTTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCCAACCAAGGGCA 540
QY 594 GCGGCGGAGTCCGGGCGCGGCCCACTGTGTGTCAGACGCGCGCGAGAGGGGCGCGAC 653
Db 541 GCGGCGGAGTCCGGGCGCGGCCCACTGTGTGTCAGACGCGCGCGAGAGGGGCGCGAC 600
QY 654 CGGACCCAGACAGAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGAAGCTTGA 713
Db 601 CGGACCCAGACAGAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGAAGCTTGA 660
QY 714 CATCCGCTCCGGGTTGAGCAGAGTACTGAGCATGTGGCAGCTTTGAGAGCGGGCGT 773
Db 661 CATCCGCTCCGGGTTGAGCAGAGTACTGAGCATGTGGCAGCTTTGAGAGCGGGCGT 720
QY 774 GGCATCCCGGCGGCCCAAGCGCTGTGGCGGCACTGACGTGTTTGGAGCGCCACCGC 833

Db 721 GGCATCCCGGCGGCCCAAGCGCGCTGTGGCGGCACTGTTTGGAGCGCCACCGC 780
QY 834 AGTCTGCGCTCAGAGGACCTGGGCTCTGTGTGTTGTGACATCAAGTTCTAGAGCTTC 893
Db 781 AGTCTGCGCTCAGAGGACCTGGGCTCTGTGTGTTGTGACATCAAGTTCTAGAGCTTC 840
QY 894 CTATCTGGAAGCTCTTGGGCGAGCTACCTGAGTGGCGCTGTGCTGAGGCGCGCGGG 953
Db 841 CTATCTGGAAGCTCTTGGGCGAGCTACCTGAGTGGCGCTGTGCTGAGGCGCGCGGG 900
QY 954 CGTGTCTGACTGAGGCGCTTGCAGAGAGCTGTGGCGCGGAGCTGTTCCCTGCTGGT 1013
Db 901 CGTGTCTGACTGAGGCGCTTGCAGAGAGCTGTGGCGCGGAGCTGTTCCCTGCTGGT 960
QY 1014 CAGTGTGATGAGGCTGACTATGAGGCTGTGGCGCGCGCTGTGCTGATGAGAGAGA 1073
Db 961 CAGTGTGATGAGGCTGACTATGAGGCTGTGGCGCGCGCTGTGCTGATGAGAGAGA 1020
QY 1074 AGGGGCGCGGCGCGCAGAGGCGCTCTGATCCAGAGCTGGCAGAGATTGATCCACCTC 1133
Db 1021 AGGGGCGCGGCGCGCAGAGGCGCTCTGATCCAGAGCTGGCAGAGATTGATCCACCTC 1080
QY 1134 CAAGTCTCCGGGCGACCTTCTCTGGAGAGAGCAATCTCTA----- 1176
Db 1081 CAAGTCTCCGGGCGACCTTCTCTGGAGAGAGCAATCTCTA----- 1140
QY 1177 -----CCCTTGACAGCCCTCC 1194
Db 1141 CTTAGCATCTTTAGAGACTGCGACAGAGCCGAGACAGAGCCCTTGAACAGCCCTCC 1200
QY 1195 CACAGAGTGTGGGCTTAGAGGCTTAACCAATTGACAGTGAAGTTCTCTCCAGACTCT 1254
Db 1201 CACAGAGTGTGGGCTTAGAGGCTTAACCAATTGACAGTGAAGTTCTCTCCAGACTCT 1260
QY 1255 CTTACCCCAAGTGTGCCCCCTTAGCCTTCGAGAGCGGGGCTGGGCTGTATCTCAAA 1314
Db 1261 CTTACCCCAAGTGTGCCCCCTTAGCCTTCGAGAGCGGGGCTGGGCTGTATCTCAAA 1320
QY 1315 GGGAGGGGCAACGCTACACACTCAACAAAGGCCCCCTGACATTTATCTCTGATCTTG 1374
Db 1321 GGGAGGGGCAACGCTACACACTCAACAAAGGCCCCCTGACATTTATCTCTGATCTTG 1380
QY 1375 GGTGTCTGACAGTGTCAAGGTGACACATGCTGTCACTAGCCCTGCTGAGA 1434
Db 1381 GGTGTCTGACAGTGTCAAGGTGACACATGCTGTCACTAGCCCTGCTGAGA 1440
QY 1435 TCTTCCCTGGGCTCTGACCTGTGCTGCTTCCAGCACAACCTTCTTGGCTTAAGGCT 1494
Db 1441 TCTTCCCTGGGCTCTGACCTGTGCTGCTTCCAGCACAACCTTCTTGGCTTAAGGCT 1500
QY 1495 TCTTCTCAGAGACTCTTAATTTTGAACCAACCAACCTGGGCTTGAACCAATCAGTGGGC 1554
Db 1501 TCTTCTCAGAGACTCTTAATTTTGAACCAACCAACCTGGGCTTGAACCAATCAGTGGGC 1560
QY 1555 ACTGAGAGTGGGGTGCACATGGGGCTGTCACTTCCAGCATCTCAGAGGAGCAG 1614
Db 1561 ACTGAGAGTGGGGTGCACATGGGGCTGTCACTTCCAGCATCTCAGAGGAGCAG 1620
QY 1615 GGCCTGCGCGAGCTTCAATTTTGAAGACCTGACTCTCTCACTTCCCTGCTGTCAAG 1674
Db 1621 GGCCTGCGCGAGCTTCAATTTTGAAGACCTGACTCTCTCACTTCCCTGCTGTCAAG 1680
QY 1675 AGCTGAACATAGACTTGAATTTGACCTGAGTGTCACTGAGTGTCACTGAGAGTGTATGCA 1734
Db 1681 AGCTGAACATAGACTTGAATTTGACCTGAGTGTCACTGAGTGTCACTGAGAGTGTATGCA 1740
QY 1735 GCATCATACCAAGGCTTACTGTTTGACATGTGGGCAAAACAGTAAACAGCACTTCTT 1794
Db 1741 GCATCATACCAAGGCTTACTGTTTGACATGTGGGCAAAACAGTAAACAGCACTTCTT 1800
QY 1795 GAAAGGGAATGCAAGGCTTTGGGGGTGATGGAAGAAAGACTTTTACAAATGATACCAAT 1854

Db 1801 GGAAAGGATGCAAGGCTTTGGGGGTGATGAGAAAGACCTTTTACAAATATACCAAT 1860

RESULT 3
CR625070 1894 bp mRNA linear HNC 21-JUL-2004
LOCUS full-length cDNA clone CS0DC001YC18 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
ACCESSION CR625070
VERSION CR625070.1 GI:50505877
KEYWORDS HNC: CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1894)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1894)
REFERENCE Direct Submission
AUTHORS Genoscope.
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source 1. 1894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC001YC18"
/issue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 91.2%; Score 1717.4; DB 3; Length 1894;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1790; Conservative 0; Mismatches 6; Indels 59; Gaps 1;
QY 75 GGCAGAAATCTGCGCGGCTCTGAGCTTGTCCGCCCTCCCTCCCGGGAATGCGCTATC 134
Db 40 GGCACCTTTGGCCGGGCTCTGAGCTTGTCCGCCCTCCCTCCCGGGAATGCGCTATC 99
QY 135 CCGGTCGACCCCGCGCCCGCTGCTGGAGAGAGATGAGTCTGAGACTAGTACGAGATGCT 194
Db 100 CCGGTCGACCCCGCGCCCGCTGCTGGAGAGAGATGAGTCTGAGACTAGTACGAGATGCT 159
QY 195 GTGCTTCAACCGTATGTGAGGTGTGGCGGGCAATGACCGAGTGGAGCTGAGCT 254
Db 160 GTGCTTCAACCGTATGTGAGGTGTGGCGGGCAATGACCGAGTGGAGCTGAGCT 219
QY 255 CCTGGCTTTTCTGCTGAGTGAAGCTCTGCGCGCGCGGAGGCTTAGCCCGGCGCGCAG 314
Db 220 CCTGGCTTTTCTGCTGAGTGAAGCTCTGCGCGCGCGGAGGCTTAGCCCGGCGCGCAG 279
QY 315 CCGGCTAGAGCTCTCTGAGAGTGGAGCGCCGCGGAGTGGCGGAGAGCAACTGCG 374
Db 280 CCGGCTAGAGCTCTCTGAGAGTGGAGCGCCGCGGAGTGGCGGAGAGCAACTGCG 339
QY 375 GCTGCTGGGGCAACTCTCTGCGCGTGTGGCCCGCCAGACGACTGCTCCGCACTGGCGCG 434
Db 340 GCTGCTGGGGCAACTCTCTGCGCGTGTGGCCCGCCAGACGACTGCTCCGCACTGGCGCG 399
QY 435 CAAGCGCGCGCGCGAGTGTCTTCAGACGCTATAGCTATGAGCACTCCAGCTCTTCAAA 494

Db 400 CAAGCGCGCGCGCGCAAGTGTCTTCAGAAAGCTATAGCTATGAGCACTCCAGCTCTTCAA 459
QY 495 GAGGACAGAGGGTAGCTGCGCGTGCCTGTGGAGTCAAGACAGTTCTGCAAAATTTTCAGCA 554
Db 460 GAGGACAGAGGGTAGCTGCGCGTGCCTGTGGAGTCAAGACAGTTCTGCAAAATTTTCAGCA 519
QY 555 GGGTCAGTGGAGAGCAAGGCTCCCCCACAAGGCGGAGCGCGGAGTGGCGCGCGCG 614
Db 520 GGGTCAGTGGAGAGCAAGGCTCCCCCACAAGGCGGAGCGCGGAGTGGCGCGCGCG 579
QY 615 CAGTGGTGTGTCAGACGCGCGCGGAGAGAGGGGCGCCAGCGCGGACCCAGCAGACAGTCA 674
Db 580 CAGTGGTGTGTCAGACGCGCGCGGAGAGAGGGGCGCCAGCGCGGACCCAGCAGACAGTCA 639
QY 675 GCCCGCAGACCTTCTCTTGAAGGAAAGTGACTGTGACATCCGAGCTCCGGGTTTCAGC 734
Db 640 GCCCGCAGACCTTCTCTTGAAGGAAAGTGACTGTGACATCCGAGCTCCGGGTTTCAGC 699
QY 735 AAGATTAATGCGAGCATGGGCGCAGCTTGGAGAGAGGCGGTGGCATCCCGGCGCGCGCG 794
Db 700 AAGATTAATGCGAGCATGGGCGCAGCTTGGAGAGAGGCGGTGGCATCCCGGCGCGCGCG 759
QY 795 GCTGGCGCGCGCAGCTGGAAGTGTGGGAGGCGCACCGGAGTGTGCGCTCAAGGGACCT 854
Db 760 GCTGGCGCGCGCAGCTGGAAGTGTGGGAGGCGCACCGGAGTGTGCGCTCAAGGGACCT 819
QY 855 GGGCTCTGTGTGTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGACGCTTTCGGG 914
Db 820 GGGCTCTGTGTGTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGACGCTTTCGGG 879
QY 915 CGACTTAATGAGTGGCGCGCGCTCTGAGAGGCGCTGGCGGGCGGTTCCTGACAGAGCGCT 974
Db 880 CGACTTAATGAGTGGCGCGCGCTCTGAGAGGCGCTGGCGGGCGGTTCCTGACAGAGCGCT 939
QY 975 GCGAGAGCTGTGGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
Db 940 GCGAGAGCTGTGGCGCGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 999
QY 1035 TGAGGCTGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1094
Db 1000 TGAGGCTGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1059
QY 1095 GGCCTCTGATTCAGAGTCTGGAGAGTATTCACCTCCAAAGTCTCCCGGCGCACCTTCT 1154
Db 1060 GGCCTCTGATTCAGAGTCTGGAGAGTATTCACCTCCAAAGTCTCCCGGCGCACCTTCT 1119
QY 1155 CTTGGAGAGAGCAACCATCTCTA----- 1176
Db 1120 CTTGGAGAGAGCAACCATCTCTAACCCTTAGAGAGTCACTAGCATCTTTGAGAGACTG 1179
QY 1177 -----CCCTTGAACAGCCCTCTCCACAGAGATGTGGCTGTGAG 1215
Db 1180 CGACAGAGACCGGAGACAGCAGGCGCCTTGAACAGCCCTCCACAGAGATGTGGCTGTGAG 1239
QY 1216 CCTAAACATTTTCAGACTAGTTCCTTCCAGACCTCTCTCAACCCCGAGGTGTGGCGCG 1275
Db 1240 CCTAAACATTTTCAGACTAGTTCCTTCCAGACCTCTCTCAACCCCGAGGTGTGGCGCG 1299
QY 1276 TTAGCTCTCGAGAGGCGGGGCTGGGCTGTATCTCGAAGGAGGAGGCGACACTACAC 1335
Db 1300 TTAGCTCTCGAGAGGCGGGGCTGGGCTGTATCTCGAAGGAGGAGGCGACACTACAC 1359
QY 1336 TCACCAAGAGCCCGCTGCAATTGATCTGATCTTGAGGTGTGCTGCACTGTCAAG 1395
Db 1360 TCACCAAGAGCCCGCTGCAATTGATCTGATCTTGAGGTGTGCTGCACTGTCAAG 1419
QY 1396 TGCACACATTCGCTATGTCTACACAGTGCCTCTGTGAGATCTTCTCTGGGCTCTGCGCT 1455
Db 1420 TGCACACATTCGCTATGTCTACACAGTGCCTCTGTGAGATCTTCTCTGGGCTCTGCGCT 1479
QY 1456 GGCCTGCTTCCAGAGCAACATCTTTGGCTTAAGGGCTTCTCTCAGAGACCTTAATT 1515
Db 1480 GGCCTGCTTCCAGAGCAACATCTTTGGCTTAAGGGCTTCTCTCAGAGACCTTAATT 1539

QY 1516 TGACCAACCAACCTGCGCTTTCAGCCATCAGTGGGCACTGAGCTGGGGTGCACATG 1575
DB 1540 TGACCAACCAACCTGCGCTTTCAGCCATCAGTGGGCACTGAGCTGGGGTGCACATG 1599
QY 1576 GGGGCTCTCACTTGGCCCAACATCTCCAGCCAGCCAGGGCCCTGGCCAGCTTCAATTT 1635
DB 1600 GGGGCTCTCACTTGGCCCAACATCTCCAGCCAGCCAGGGCCCTGGCCAGCTTCAATTT 1659
QY 1636 ACAGACCTGACTCTCTCACTTCCCTGCTGCTGCAAGCTGACATGACTTGGACT 1695
DB 1660 ACAGACCTGACTCTCTCACTTCCCTGCTGCTGCTGCAAGCTGACATGACTTGGACT 1719
QY 1696 TGAATGTCACCTGAGAGTGTCAATGAGAGTGTATGAGCAGCATCAACAAGCTTACTG 1755
DB 1720 TGAATGTCACCTGAGAGTGTCAATGAGAGTGTATGAGCAGCATCAACAAGCTTACTG 1779
QY 1756 TTGACATGAGGGCCCAACCAAGTAAACAGCCACTTCTTGAAGAGGATGCAAGGCTT 1815
DB 1780 TTGACATGAGGGCCCAACCAAGTAAACAGCCACTTCTTGAAGAGGATGCAAGGCTT 1839
QY 1816 TGGGGGATGAGAAAGACCTTTTACAATGATACCAATTAACTGCGCTGGAAA 1870
DB 1840 TGGGGGATGAGAAAGACCTTTTACAATGATACCAATTAACTGCGCTGGAAA 1894

RESULT 4
CR599685
LOCUS
DEFINITION
full-length cDNA clone CSOD1069YH17 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR599685
VERSION
CR599685.1 GI:50480492
KEYWORDS
HTC; CNSLT cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1510)
L.I.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1510)
REFERENCE
AUTHORS
JOURNAL
REMARK

COMMENT
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1.1510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1069YH17"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 71.9%; Score 1354.4; DB 3; Length 1510;
Best Local Similarity 96.0%; Pred. No. 3.8e-313;
Matches 1424; Conservative 0; Mismatches 1; Indels 59; Gaps 1;
QY 449 CAGGTCTCCAGAACGCTATAGCTATGAGCACTCCAGCTCTTCAAAAGAGCAAGAGGTA 508
DB 27 CTGTGTCTCAGAACGCTATAGCTATGAGCACTCCAGCTCTTCAAAAGAGCAAGAGGTA 86

QY 509 GCTCCCTGCGCGTGGCACTCAAGCAAGTTCGTGCAATTCTCAGCAGAGGTCAAGTGGAGA 568
DB 87 GCTCCCTGCGCGTGGCACTCAAGCAAGTTCGTGCAATTCTCAGCAGAGGTCAAGTGGAGA 146
QY 569 CAGGCTCCCCCAACCAAGCGGAGCGGAGAGTCCGGGCGGGCCAGTGTGTGTCCA 628
DB 147 CAGGCTCCCCCAACCAAGCGGAGCGGAGAGTCCGGGCGGGCCAGTGTGTGTCCA 206
QY 629 GAGGGGGGAGAGAGGGGGCCCAAGCGGCAAGCCCAAGCAAGTCAAGAGCCCGGAGACTT 688
DB 207 GAGGGGGGAGAGAGGGGGCCCAAGCGGCAAGCCCAAGCAAGTCAAGAGCCCGGAGACTT 266
QY 689 CCTCTGAAGCAAGTGAAGTCAATCCGAGTCCGGGTTCCAGCAGAGTACTGCCAGC 748
DB 267 CCTCTGAAGCAAGTGAAGTCAATCCGAGTCCGGGTTCCAGCAGAGTACTGCCAGC 326
QY 749 ATGGGCGAGCTTGGAGCAGAGGCGTGGCATCCCGGCGGCCCAAGCGCTGGCGGAGC 808
DB 327 ATGGGCGAGCTTGGAGCAGAGGCGTGGCATCCCGGCGGCCCAAGCGCTGGCGGAGC 386
QY 809 TGAAGGTGTTGGGCAAGGCGCAGCGAGTGTGGCTCAAGGGACTGGGGCTGTGTTT 868
DB 387 TGAAGGTGTTGGGCAAGGCGCAGCGAGTGTGGCTCAAGGGACTGGGGCTGTGTTT 446
QY 869 GTGACATCAAGTTCAGAGCTCTCTATCTGAGACGCTTCTGGGGCGACTGAGATG 928
DB 447 GTGACATCAAGTTCAGAGCTCTCTATCTGAGACGCTTCTGGGGCGACTGAGATG 506
QY 929 GCGGCTGTGTCAGAGGCGCTGCGGGCGGTGTTCTGACTAGGCGCTGCGAGAGCTGTGG 988
DB 507 GCGGCTGTGTCAGAGGCGCTGCGGGCGGTGTTCTGACTAGGCGCTGCGAGAGCTGTGG 566
QY 989 GCGGAGGCGTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
DB 567 GCGGAGGCGTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 626
QY 1049 GCGGCTGTGTCAGAGGAG 1108
DB 627 GCGGCTGTGTCAGAGGAG 686
QY 1109 GAGCTGAGAGATGATGATCCAGCTCCAGAGTCTCCGGGCACTTCTCTGGAGAGAGAC 1168
DB 687 GAGCTGAGAGATGATGATCCAGCTCCAGAGTCTCCGGGCACTTCTCTGGAGAGAGAC 746
QY 1169 GAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1176
DB 747 GAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 806
QY 1177 -----CCCTTGAAGAGCGCTCCCAAGAGATGAGGCTTGAAGGCTTAACATTTCC 1229
DB 807 CAGCAGGCGCTTGAAGAGCGCTCCCAAGAGATGAGGCTTGAAGGCTTAACATTTCC 866
QY 1230 AGCTGAGTTCCTTCCAGAGCTCTCTCTCAAGCCCAAGTGTGCCCTTGAAGCTTCCGAGG 1289
DB 867 AGCTGAGTTCCTTCCAGAGCTCTCTCTCAAGCCCAAGTGTGCCCTTGAAGCTTCCGAGG 926
QY 1290 CCGGGGCTGGGCTGTATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1349
DB 927 CCGGGGCTGGGCTGTATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 986
QY 1350 CCTGACATTTGATCTGATCTTGGGCTGTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAG 1409
DB 987 CCTGACATTTGATCTGATCTTGGGCTGTGAGCTGTGAGAGAGAGAGAGAGAGAGAGAG 1046
QY 1410 CATGCTCACTGCGCTGCTGAGATCTTCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTG 1469
DB 1047 CATGCTCACTGCGCTGCTGAGATCTTCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTG 1106
QY 1470 CAGACATTTTGGCTAAGAGGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
DB 1107 CAGACATTTTGGCTAAGAGGCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166

QY 1530 CTGGGCTTACGCCACATCACTAGTGGGCACTGAGCTGGGGTGCACATGAGGGCTTCTCACT 1589
DB 1167 CTGGGGCTTACGCCACATCACTAGTGGGCACTGAGCTGGGGTGCACATGAGGGCTTCTCACT 1226
QY 1590 TGGCCACACATCTCCAGCCGAGCCGAGGGCCCTGGCCAGCTTCAATTTCACAGACCTGACTCT 1649
DB 1227 TGGCCACACATCTCCAGCCGAGCCGAGGGCCCTGGCCAGCTTCAATTTCACAGACCTGACTCT 1286
QY 1650 CCTGACCTTCCCTCCCTGCTGTCCAGAGCTGAACATGACTTGCACCTTGCATCTGG 1709
DB 1287 CCTGACCTTCCCTCCCTGCTGTCCAGAGCTGAACATGACTTGCATCTGATGCACTGG 1346
QY 1710 AGGTGACATGAGAGTGTATGCGACATCATACCAAGGCTTCTGTGACATGAGGGCC 1769
DB 1347 AGGTGACATGAGAGTGTATGCGACATCATACCAAGGCTTCTGTGACATGAGGGCC 1406
QY 1770 AAAACGATTAACAGCACCTTCTTGTGAAAAGGAATGCAAGGCTTGTGGGGTGTATGAA 1829
DB 1407 AAAACGATTAACAGCACCTTCTTGTGAAAAGGAATGCAAGGCTTGTGGGGTGTATGAA 1466
QY 1830 AAGACCTTTTACAAATGATACCAATTAACTGCCCTGAAAAGG 1873
DB 1467 AAGACCTTTTACAAATGATACCAATTAACTGCCCTGAAAAGG 1510

RESULT 5
CRS90827 1586 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DE001YM1 of Placenta of Homo sapiens
DEFINITION (human).

ACCESSION CRS90827
VERSION CRS90827.1 GI:50471634
KEYWORDS HTC; CDSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1586)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polyes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Peng Liang Email: liang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue

REFERENCE 2 (bases 1 to 1586)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage: BP 131 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr

COMMENT - Web: www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.

FEATURES
source Location/Qualifiers
1..1586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE001YM1"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 71.6%; Score 1348.8; DB 3; Length 1586;
Best Local Similarity 94.6%; Pred.No.8.5e-312; Indels 59; Gaps 1;
Matches 1431; Conservative 0; Mismatches 22;

QY 415 CTGCTCCGACCTGGCGCGCAAGCGCGCCGCAAGTGTCTCCAGAACCTATAGCTAT 474
DB 75 CGGCGCTCGCCAGAGACTATCGAGCCGCAAGATCTGTCTCCAGAACCTATAGCTAT 134

QY 475 GGACCTCCAGCTCTTCAAAAGAGACAGAGGGTACTGCGCTGCGCTGGGACGTCAAGC 534
DB 135 GGACCTCCAGCTCTTCAAAAGAGAGACAGAGGGTACTGCGCTGCGCTGGGACGTCAAGC 194
QY 535 AGTTCGCAAAATTTCTCAGAGAGGTCACTGGAGAGACAGGCTCCCGCCCAACCAAGCGGACG 594
DB 195 AGTTCGCAAAATTTCTCAGAGAGGTCACTGGAGAGAGAGGCTCCCGCCCAACCAAGCGGACG 254
QY 595 CGGCGAGTGGGGCGGCGCCAGTGGTGGTGCAGAGCGGCGGAGAGAGGGCCCGACGC 654
DB 255 CGGCGAGTGGGGCGGCGCCAGTGGTGGTGCAGAGCGGCGGAGAGAGGGCCCGACGC 314
QY 655 GCAACCCAGACAGTCAAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAACTGACTGTGAC 714
DB 315 GCAACCCAGACAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAACTGACTGTGAC 374
QY 715 ATCCGGCTCCGGGTTTGAAGAGTACTGAGCATATGGCCAGCCTTGGAGAGAGGCGTG 774
DB 375 ATCCGGCTCCGGGTTTGAAGAGTACTGAGCATATGGCCAGCCTTGGAGAGAGGCGTG 434
QY 775 GCATCCCGGCGGCGCCAGGCGCTGGCGCGGCAAGCTGATTTGGGAGGCGCACGCA 834
DB 435 GCATCCCGGCGGCGCCAGGCGCTGGCGCGGCAAGCTGATTTGGGAGGCGCACGCA 494
QY 835 GTGCTGCGCTCAAGGAGACCTGGGCTCTGTGTGTGTGTGATGATCAAGTCTCAAGCTTCC 894
DB 495 GTGCTGCGCTCAAGGAGACCTGGGCTCTGTGTGTGTGTGATGATCAAGTCTCAAGCTTCC 554
QY 895 TATCTGAGAGCCTTCTGGGGGCACTACTGAGTGGGCGCCTGCTGAGAGCCCTGCGGGGC 954
DB 555 TATCTGAGAGCCTTCTGGGGGCACTACTGAGTGGGCGCCTGCTGAGAGCCCTGCGGGGC 614
QY 955 GTGTTCTGATGAGGCGCTGCGGAGAGGCTGTGGGCGGAGAGGCTGTGCTGTGTC 1014
DB 615 GTGTTCTGATGAGGCGCTGCGGAGAGGCTGTGGGCGGAGAGGCTGTGCTGTGTC 674
QY 1015 AGTGTGAGTGAAGAGCTGATATAGAGTGGCGCGGCGGCGCTGTGTGTATGAGAGAGGAA 1074
DB 675 AGTGTGAGTGAAGAGCTGATATAGAGTGGCGCGGCGGCGCTGTGTGTATGAGAGAGGAA 734
QY 1075 GGGGGGCGGCGCGCCAGAGAGGCTCTGATCCAGAGCTGGAGATTAATCCACCTCC 1134
DB 735 GGGGGGCGGCGCGCCAGAGAGGCTCTGATCCAGAGCTGGAGATTAATCCACCTCC 794
QY 1135 AAGTCTCGGGCGACCTTCTCTGGAGAGACGACATCTCTA----- 1176
DB 795 AAGTCTCGGGCGACCTTCTCTGGAGAGACGACATCTCTA----- 854
QY 1177 -----CCCTTGAACAGCCCTCCG 1195
DB 855 TCTAGATCTTGAAGAGCTGCGACAGAGCCGGGACAGAGCCCTTGAACAGCCCTCCG 914
QY 1196 ACAGAGTGTGGGCTTGAGAGGCTTAACCATTTTCAGACTAGTTTCTTCCAGACTCTTC 1255
DB 915 ACAGAGTGTGGGCTTGAGAGGCTTAACCATTTTCAGACTAGTTTCTTCCAGACTCTTC 974
QY 1256 CTACCCCAAGTGTGCGCCCTTAGCCTCGGAGGCGGGGCTGGGCTGTATCTCAAG 1315
DB 975 CTACCCCAAGTGTGCGCCCTTAGCCTCGGAGGCGGGGCTGGGCTGTATCTCAAG 1034
QY 1316 GAGAGGGGACAGTACACTACCAAAAGCCCTCTGACATTTGATATCTGATCTTGG 1375
DB 1035 GAGAGGGGACAGTACACTACCAAAAGCCCTCTGACATTTGATATCTGATCTTGG 1094
QY 1376 GCTGTGACATGTCAAGAGTGCACACTCGCTCATGCTCACTGAGCCCTGTAGAT 1435
DB 1095 GCTGTGACATGTCAAGAGTGCACACTCGCTCATGCTCACTGAGCCCTGTAGAT 1154
QY 1436 CTTCCTTGAGGCTCTGCGCTGCGCTTTCAGACACACTTCTTTGGCTTAAGGCGTT 1495
DB 1155 CTTCCTTGAGGCTCTGCGCTGCGCTTTCAGACACACTTCTTTGGCTTAAGGCGTT 1214
QY 1496 CTCTCTCAGGACCTTAATTGACCAACCAACCTGAGCTTACGCAATCAGTGGCA 1555

Db	Accession	Length	Score	DB 3	Length	DB 1	DB 2	DB 4	DB 5	DB 6	DB 7	DB 8	DB 9	DB 10	DB 11	DB 12	DB 13	DB 14	DB 15	DB 16	DB 17	DB 18	DB 19	DB 20	DB 21	DB 22	DB 23	DB 24	DB 25	DB 26	DB 27	DB 28	DB 29	DB 30	DB 31	DB 32	DB 33	DB 34	DB 35	DB 36	DB 37	DB 38	DB 39	DB 40	DB 41	DB 42	DB 43	DB 44	DB 45	DB 46	DB 47	DB 48	DB 49	DB 50	DB 51	DB 52	DB 53	DB 54	DB 55	DB 56	DB 57	DB 58	DB 59	DB 60	DB 61	DB 62	DB 63	DB 64	DB 65	DB 66	DB 67	DB 68	DB 69	DB 70	DB 71	DB 72	DB 73	DB 74	DB 75	DB 76	DB 77	DB 78	DB 79	DB 80	DB 81	DB 82	DB 83	DB 84	DB 85	DB 86	DB 87	DB 88	DB 89	DB 90	DB 91	DB 92	DB 93	DB 94	DB 95	DB 96	DB 97	DB 98	DB 99	DB 100	DB 101	DB 102	DB 103	DB 104	DB 105	DB 106	DB 107	DB 108	DB 109	DB 110	DB 111	DB 112	DB 113	DB 114	DB 115	DB 116	DB 117	DB 118	DB 119	DB 120	DB 121	DB 122	DB 123	DB 124	DB 125	DB 126	DB 127	DB 128	DB 129	DB 130	DB 131	DB 132	DB 133	DB 134	DB 135	DB 136	DB 137	DB 138	DB 139	DB 140	DB 141	DB 142	DB 143	DB 144	DB 145	DB 146	DB 147	DB 148	DB 149	DB 150	DB 151	DB 152	DB 153	DB 154	DB 155	DB 156	DB 157	DB 158	DB 159	DB 160	DB 161	DB 162	DB 163	DB 164	DB 165	DB 166	DB 167	DB 168	DB 169	DB 170	DB 171	DB 172	DB 173	DB 174	DB 175	DB 176	DB 177	DB 178	DB 179	DB 180	DB 181	DB 182	DB 183	DB 184	DB 185	DB 186	DB 187	DB 188	DB 189	DB 190	DB 191	DB 192	DB 193	DB 194	DB 195	DB 196	DB 197	DB 198	DB 199	DB 200	DB 201	DB 202	DB 203	DB 204	DB 205	DB 206	DB 207	DB 208	DB 209	DB 210	DB 211	DB 212	DB 213	DB 214	DB 215	DB 216	DB 217	DB 218	DB 219	DB 220	DB 221	DB 222	DB 223	DB 224	DB 225	DB 226	DB 227	DB 228	DB 229	DB 230	DB 231	DB 232	DB 233	DB 234	DB 235	DB 236	DB 237	DB 238	DB 239	DB 240	DB 241	DB 242	DB 243	DB 244	DB 245	DB 246	DB 247	DB 248	DB 249	DB 250	DB 251	DB 252	DB 253	DB 254	DB 255	DB 256	DB 257	DB 258	DB 259	DB 260	DB 261	DB 262	DB 263	DB 264	DB 265	DB 266	DB 267	DB 268	DB 269	DB 270	DB 271	DB 272	DB 273	DB 274	DB 275	DB 276	DB 277	DB 278	DB 279	DB 280	DB 281	DB 282	DB 283	DB 284	DB 285	DB 286	DB 287	DB 288	DB 289	DB 290	DB 291	DB 292	DB 293	DB 294	DB 295	DB 296	DB 297	DB 298	DB 299	DB 300	DB 301	DB 302	DB 303	DB 304	DB 305	DB 306	DB 307	DB 308	DB 309	DB 310	DB 311	DB 312	DB 313	DB 314	DB 315	DB 316	DB 317	DB 318	DB 319	DB 320	DB 321	DB 322	DB 323	DB 324	DB 325	DB 326	DB 327	DB 328	DB 329	DB 330	DB 331	DB 332	DB 333	DB 334	DB 335	DB 336	DB 337	DB 338	DB 339	DB 340	DB 341	DB 342	DB 343	DB 344	DB 345	DB 346	DB 347	DB 348	DB 349	DB 350	DB 351	DB 352	DB 353	DB 354	DB 355	DB 356	DB 357	DB 358	DB 359	DB 360	DB 361	DB 362	DB 363	DB 364	DB 365	DB 366	DB 367	DB 368	DB 369	DB 370	DB 371	DB 372	DB 373	DB 374	DB 375	DB 376	DB 377	DB 378
----	-----------	--------	-------	------	--------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

Db	1	GAAGAATCCAAACAAAGTGGCTGGCGGCGCTGGCCCAAGATGATCGAATCCGCAAAATCTG	60
OY	86	GCCGGGTTTCTGAGCTTTTCGACCTCCCTCCCTCCCGGAAATGCGCTATCCGGATCGAACCC	145
Db	61	GCCGGGTTTCTGAGCTTTGTTCCGCTCCCTCCCTCCCGGAAATGGCGCTATCCGGATCGAACCC	120
OY	146	CGGCCCCGTGTGGGAAGAGATGATGTCCTGCACTTACGGGAATGTGTCTGCTTCAAC	205
Db	121	CGGCCCCGTGTGGGAAGAGATGATGTCCTGCACTTACGGGAATGTGTCTGCTTCAAC	180
OY	206	GTATGTTGAGGTGTGGGCGGGCAATGACCGAGTGCAGTGTGAGCTCCCTGGCTTTTC	265
Db	181	GTATGTTGAGGTGTGGGCGGGCAATGACCGAGTGCAGTGTGAGCTCCCTGGCTTTTC	240
OY	266	TGCTGGAATGAGAGCTCTTGCGCGCCGCGAGGCTTATGACCCTGGGCCCGGACGCGCTTAGAGC	325
Db	241	TGCTGGAATGAGAGCTCTTGCGCGCCGCGAGGCTTATGACCCTGGGCCCGGACGCGCTTAGAGC	300
OY	326	TCTGTCCTGGAAGTGGAGCGCCGCGGAGTGGCGGAGAGCACTTGCGCTGCTGGGCGC	385
Db	301	TCTGTCCTGGAAGTGGAGCGCCGCGGAGTGGCGGAGAGCACTTGCGCTGCTGGGCGC	360
OY	386	AACCTCCGCGGTGTGTGTGGGCCCCGCAAGACCTGTCTGCGCGACCTTGCGCGCGCAAGGGGCGCC	445
Db	361	AACCTCCGCGGTGTGTGTGGGCCCCGCAAGACCTGTCTGCGCGACCTTGCGCGCGCAAGGGGCGCC	420
OY	446	GGCAGAGTGTCCAGAAAGCTATATGATATGAGCACTTCAGCTCTTCAAAGAGACAGAGG	505
Db	421	GGCAGAGTGTCTCCAGAAAGCTATATGATATGAGCACTTCAGCTCTTCAAAGAGACAGAGG	480
OY	506	GTAAGCTCCGCTCGCGGTGCGGAGTCAAGCAATTTCTGAGCAAGGTCAGTGGG	565
Db	481	GTAAGCTCCGCTCGCGGTGCGGAGTCAAGCAATTTCTGAGCAAGGTCAGTGGG	540
OY	566	AGAAAGGCTCCCCCCCCCAACCAAGGCGAGCGGCGGAGTCCGGGCGCGGCCCAATGTGTGTG	625
Db	541	AGAAAGGCTCCCCCCCCCAACCAAGGCGAGCGGCGGAGTCCGGGCGCGGCCCAATGTGTGTG	600
OY	626	CCAGACGCGCGGCGAGAGGGGCCCCAGCCGCAACCCCAAGACAGTCAAGAGCCCGCCAGAC	685
Db	601	CCAGACGCGCGGCGAGAGGGGCCCCAGCCGCAACCCCAAGACAGTCAAGAGCCCGCCAGAC	660
OY	686	CTTCTCTGAAAGGCAAGTGAACCTGTGACATCCCGGCTCCGGGTTTGAAGCAAGTACTGCG	745
Db	661	CTTCTCTGAAAGGCAAGTGAACCTGTGACATCCCGGCTCCGGGTTTGAAGCAAGTACTGCG	720
OY	746	AGCATGGGACGCTTGGAGAGAGGCGTGGCATTCGCGCGGCCCCAGAGCGCTGGCGCGGC	805
Db	721	AGCATGGGACGCTTGGAGAGAGGCGTGGCATTCGCGCGGCCCCAGAGCGCTGGCGCGGC	780
OY	806	AGCTGGAAGGTGTTGGGCGAGGCAACGCGAGTGGTGGGCTCAAGGGAACTGGGGCTGTGGG	865
Db	781	AGCTGGAAGGTGTTGGGCGAGGCAACGCGAGTGGTGGGCTCAAGGGAACTGGGGCTGTGGG	840
OY	866	TTTGTGACATCAAGTTTCTCAGAGCTCTCTTATCTGACAGGCTTCTGGGCGCACTTACCTGA	925
Db	841	TTTGTGACATCAAGTTTCTCAGAGCTCTCTTATCTGACAGGCTTCTGGGCGCACTTACCTGA	900
OY	926	GTGGCGCCCTGCTGCAAGGCCCCGTGGGGCGTTTCTGACTGAGGCCCTTGAGAGGCTG	985
Db	901	GTGGCGCCCTGCTGCAAGGCCCCGTGGGGCGTTTCTGACTGAGGCCCTTGAGAGGCTG	960
OY	986	TGGGCGCGGAGAGGCTTTCGCTGCTGTGATCAAGTGTGATGAGGCTGACATATGAGGCTGGCC	1045
Db	961	TGGGCGCGGAGAGGCTTTCGCTGCTGTGATCAAGTGTGATGAGGCTGACATATGAGGCTGGCC	1020
OY	1046	GAGCGCGGCTGTTGCTGATGAGAGGAGAGAGGGG 1078	
Db	1021	GAGCGCGGCTGTTGCTGATGAGAGGAGAGAGGGG 1053	

BX342837 1046 bp mRNA linear EST 07-APR-2004
 LOCUS BX342837 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DL006YL09 5-PRIME, mRNA sequence.
 ACCESSION BX342837
 VERSION BX342837.2 GI:46266793
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1046)
 Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 1, 2003 this sequence version replaced gi:30313128.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9074.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0DL006CP05QPlc=9074.r.
 FEATURES
 source
 Location/Qualifiers
 1..1046
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DL006YL09"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Query Match 50.6%; Score 952.4; DB 5; Length 1046;
 Best Local Similarity 96.7%; Pred. No. 6,5e-217;
 Matches 1018; Conservative 15; Mismatches 12; Indels 8; Gaps 6;
 QY 26 GAAAGATCCAAACAGTGGCTGGCGCGCTGCGCCAGAGTCTTCGAGCCCGCAATCTG 85
 DB 1 GAAAGATCCAAACAGTGGCTGGCGCGCTGCGCCAGAGTCTTCGAGCCCGCAATCTG 60
 QY 86 GCGGGGTTTGACCTGTTTCCGCTCCCTCCCGGGGAATGGGCTATCCGGGTTCAGACC 145
 DB 61 GCGGGGTTTGACCTGTTTCCGCTCCCTCCCGGGGAATGGGCTATCCGGGTTCAGACC 120
 QY 146 CGGCCCCGTGCTGGAGAGATGATGCTGACTACTAGCGGAGTGTCTGCTTCAACC 205
 DB 121 CGGCCCCGTGCTGGAGAGATGATGCTGACTACTAGCGGAGTGTCTGCTTCAACC 180
 QY 206 GATATGTTTCAGAGTGTGGCGCGGCAACTGACCGAGTCCGAGTCTTGGCTTTTC 265
 DB 181 GATATGTTTCAGAGTGTGGCGCGGCAACTGACCGAGTCCGAGTCTTGGCTTTTC 240
 QY 266 TGTGTGATGAGGCTCTCTGGGCGCGCGAGGCTTAAGCCCGGGCCGAGAGGGCTAGAGC 325
 DB 241 TGTGTGATGAGGCTCTCTGGGCGCGCGAGGCTTAAGCCCGGGCCGAGAGGGCTAGAGC 299
 QY 326 TCTGTGATGAGTGTGAGCGCGCGGCGAGTGTGCGGCAACTGCTGCTGCTGCTGCTG 385
 DB 300 TCTGTGATGAGTGTGAGCGCGCGGCGAGTGTGCGGCAACTGCTGCTGCTGCTGCTG 359
 QY 386 AACTCTGCGGCTGTGGCCCGCAGCACTGTGCTGCCGACCTTGGCGCGCAAGCGGCGCC 445

DB 360 AACTCTGCGGCTGTGGCCCGCAGCACTGTGCTGCCACTTGGCGCCAAAS--GCGCC 417
 QY 446 GGCAGATGCTCCAGAACCTATAGTATGAGACCTCCAGCTCTTCAAGAGAGACAGAG 505
 DB 418 GGCAGATGCTCCAGAACCTATAGTATGAGACCTCCAGCTCTTCAAGAGAGACAGAG 477
 QY 506 GATAGTCCCGTGGCGCGCTGAGTCAAGATTTCTGCAAAATTTCTAGCAGGGTCACTGGG 565
 DB 478 GATAGTCCCGTGGCGCGCTGAGTCAAGATTTCTGCAAAATTTCTAGCAGGGTCACTGGG 537
 QY 566 AGACAGGCTCCCGCCCAACCAAGCGGAGCGGAGTGTGGGCGCGCCAGTGGGTG 625
 DB 538 AGACAGGCTCCCGCCCAACCAAGCGGAGCGGAGTGTGGGCGCGCCAGTGGGTG 597
 QY 626 CCAGACGCGGCGGAGAGGGGCGCCAGCCGCAAGCCAGAGAGTCAAGAGCCCGCAGAC 685
 DB 598 CCAGACGCGGCGGAGAGGGGCGCCAGCCGCAAGCCAGAGAGTCAAGAGCCCGCAGAC 657
 QY 686 CTTCCTCTGAGGCAAAAGTACCTGTGACATCCGGCTCCGGGTTGAGCAGAGTACTGGC 745
 DB 658 CTTCCTCTGAGGCAAAAGTACCTGTGACATCCGGCTCCGGGTTGAGCAGAGTACTGGC 717
 QY 746 AGCATGGGCGAGCCTTGGAGCAGGGGCTGACATCCCGGCGGCGCCAGAGGCGCTGGCGGC 805
 DB 718 AGCATGGGCGAGCCTTGGAGCAGGGGCTGACATCCCGGCGGCGCCAGAGGCGCTGGCGGC 777
 QY 806 AGCTGACGCTGTTTGGGCGAGCCAGCCAGTGTGCTGCTCAAGAGACCTTGGCTGTGG 865
 DB 778 AGCTGACGCTGTTTGGGCGAGCCAGCCAGTGTGCTGCTCAARAGACCTTGGCTGTGG 837
 QY 866 TTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTT--GGGCGCACTACTGG 924
 DB 838 TTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTTCKGGGCGCACTACTGG 897
 QY 925 AGTGGCGCCCTGTGAGGCGCCGCGGGGCGTGTCTGACTGAGGCGCTGCGAGAGGCT 984
 DB 898 AGTGGCGCCCTGTGAGGCGCCGCGGGGCGTGTCTGACTGAGGCGCTGCGAGAGGCT 957
 QY 985 GTGGGCGGAGGCTGTGCGCTGCTGATCAAGTGTGATGAGGCTGACTATGAGGCTGGC 1044
 DB 958 GTGGGCGGAGG--GATGTTGCGCTGCTGATGAGTGTGAGG--GATGATATATAGG--TGG 1013
 QY 1045 CGGCGCGGCTGTGCTGATGAGAGAGAGG 1077
 DB 1014 CGGCGCGGCTGTGCTGATGRRRAGAGG 1046
 RESULT 8
 BX384448 1021 bp mRNA linear EST 23-APR-2004
 LOCUS BX384448 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DK010YK16 5-PRIME, mRNA sequence.
 ACCESSION BX384448
 VERSION BX384448.2 GI:46557783
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1021)
 Li, W.B., Gruber, C., Jessee, J. and Polyes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 8, 2003 this sequence version replaced gi:30440338.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster 9074.r

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?c=CS0DK010BF08QPlc=9074.r>.

FEATURES

Location/Qualifiers

```
1..1021
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="CS0DK010YK16"
  /cell_type="HELA CELLS COT 25-NORMALIZED"
  /cell_line="HELA"
  /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
  /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 47.6%; Score 896.8; DB 5; Length 1021;
Best Local Similarity 94.7%; Pred. No. 1.3e-203;
Matches 975; Conservative 16; Mismatches 28; Indels 11; Gaps 6;

```
QY 54 GTGCGCCAGAGTATCGACGCCAAGTCTGCGCGGTTCTGAGCTTCCGCTCC 113
DB 1 GTGCGCCAGAGTATCGACGCCAAGTCTGCGCGGTTCTGAGCTTCCGCTCC 60
QY 114 TCCCGCGGAGTGGCGTATCCGGGTGACCCCGCGCGCGTGGAGAGAGATGATG 173
DB 61 TCCCGCGGAGTGGCGTATCCGGGTGACCCCGCGCGCGTGGAGAGAGATGATG 120
QY 174 CTTGACTACTAGCGAGTGTCTGCTTCACTGATGATGATGATGATGATGATG 233
DB 121 CTTGACTACTAGCGAGTGTCTGCTTCACTGATGATGATGATGATGATGATG 180
QY 234 GACCGAGTGGAGTGGAGTCTTCTGCTTCTGCTGATGATGATGATGATGATG 293
DB 181 GACCGAGTGGAGTGGAGTCTTCTGCTTCTGCTGATGATGATGATGATGATG 240
QY 294 AGGCTTAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 353
DB 241 AGGCTTAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 299
QY 354 GTGCGCGGAGAGCACTGCGCGCTGCTGCGGCAACTCTCGCGCTGCGCGCGCA 413
DB 300 GTGCGCGGAGAGCACTGCGCGCTGCTGCGGCAACTCTCGCGCTGCGCGCGCA 359
QY 414 CTTGCTGCGCACTGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 473
DB 360 CTTGCTGCGCACTGCGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 417
QY 474 TGGCAGCTTCAGCTTCAAGAG-GACAGAGGTGATGCTCCGCGCGCGCGCA 532
DB 418 TGGCAGCTTCAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
QY 533 GCAGTTCTGAATTTCTCAGCAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 592
DB 478 GCAGTTCTGAATTTCTCAGCAGGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 537
QY 593 AGCGCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 652
DB 538 AGCGCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
QY 653 CGCGACCCCGAGAGTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 712
DB 598 CGCGACCCCGAGAGTGTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
QY 713 ACATCCGAGTCCGAGTGTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 772
DB 658 ACATCCGAGTCCGAGTGTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAG 717
QY 773 TGGCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 832
```

```
DB 718 TGGCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
QY 833 CAGTGTGCGCTCAGAGGACCTGAGGCTGTGAGTTGTGACATCAATCTCAGAGCTT 892
DB 778 CAGTGTGCGCTCAGAGGACCTGAGGCTGTGAGTTGTGACATCAATCTCAGAGCTT 837
QY 893 CTTATCGACCGCTTCTGCGGCGGAGTACTGAGTGGCGCGCTGCTGCTGCGCG 952
DB 838 CTTATCGACCGCTTCTGCGGCGGAGTACTGAGTGGCGCGCTGCTGCTGCGCG 896
QY 953 GCGTGTCTGACTGAGGCGCGCTGCGAGAGCTGTGAGGAGCTGTTGCTGCTG 1012
DB 897 GCGTGTCTGACTGAGGCGCGCTGCGAGAGCTGTGAGGAGCTGTTGCTGCTG 955
QY 1013 TCAGTGTGATGAGGCGTACTATGAGCTGCGCGCGCGCGCGCTGTTGATGAGAGG 1072
DB 956 TCAGTGTGATGAGGCGTACTATGAGCTGCGCGCGCGCGCGCTGTTGATGAGAGG 1010
QY 1073 AAGCGGCGCG 1082
DB 1011 AAGCGGCGCG 1020
```

RESULT 9
BX374729
LOCUS
DEFINITION
BX374729 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC001YK18 5-PRIME, mRNA sequence.
ACCESSION
BX374729
VERSION
BX374729.2 GI:46557336
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
On May 8, 2003 this sequence version replaced gi:30438519.

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Creneau, CP 5706 - 91057 Evry cedex - FRANCE
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=CS0DC001B09QPlc=9074.r>.

FEATURES

Location/Qualifiers

```
1..1091
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="CS0DC001YK18"
  /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
  /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
  /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 45.3%; Score 853.4; DB 5; Length 1091;
Best Local Similarity 94.4%; Pred. No. 3.3e-193;
Matches 954; Conservative 15; Mismatches 32; Indels 10; Gaps 8;

```
QY 75 GCCAGATCTGCGCGGCTTCTGAGCTTGTCCGCTCCCTCCCGGGAATGGCGCTATC 134
DB 40 GGCACCTTTGGCGGGTTCTGAGCTTGTCCGCTCCCTCCCGGGAATGGCGCTATC 99
```


QY	135	CGGGTCGACCCCGGCGCCCGTGTCTGGAGAGAGAGATGAGTGGCTTGACATCTCAACGGGATCT	194
Db	100	CGGGTCGACCCCGGCGCCCGTGTCTGGAGAGAGATGAGTGGCTTGACATCTCAACGGGATCT	159
QY	195	GTCGCTTCAACCTGATGTTTCGAGTGTGTGGGCGGGCACTGACCCGAGTGCAGCTTGAGACT	254
Db	160	GTCGCTTCAACCTGATGTTTCGAGTGTGTGGGCGGGCACTGACCCGAGTGCAGCTTGAGACT	219
QY	255	CTTGGGCTTTTGTGATGATGAGGCTCCGTGGGCGCGCGGAGGAGGCTTAGCGCGGGCCCGAG	314
Db	220	CTTGGGCTTTTGTGATGATGAGGCTCCGTGGGCGCGCGGAGGAGGCTTAGCGCGGGCCCGAG	278
QY	315	CGGCTTAGAGCTCTCTGTGAGCTTGAGAGCGCGCGGAGTGCGGCGAGAGCAACTTGTGCG	374
Db	279	CGGCTTAGAGCTCTCTGTGAGCTTGAGAGCGCGCGGAGTGCGGCGAGAGCAACTTGTGCG	338
QY	375	GCTGCTGGGGGCACTTCTTGCCTGGCTGTGGGCCCGGACAGACCTTGTCTGCCGACCTTGGCGG	434
Db	339	GCTGCTGGGGGCACTTCTTGCCTGGCTGTGGGCCCGGACAGACCTTGTCTGCCGACCTTGGCGG	398
QY	435	CAACGGGCGCGGCGACAGTGTCTCCAGAACGCTATAGCTATGAGCACTTCACTCTTCAAA	494
Db	399	CAA--MEGGCGGCGGACAGTGTCTCCAGAACGCTATAGCTATGAGCACTTCACTCTTCAAA	456
QY	495	GAGGACAGAGGGTACTTCCCTCGCCGTGCGGAGTCAAGCAAGTTTGTCAAAATTCTACGA	554
Db	457	GAGGACATAGGGTACTTCCCTCGCCGTGCGGAGTCAAGCAAGTTTGTCAAAATTCTACGA	516
QY	555	GGGTCAGTGGGAGACAGGCTCCCCCGCCCAACAGAGGAGGCGGCGGAGTCCGGGGCCGGCG	614
Db	517	GGGTCAGTGGGAGACAGGCTCCCCCGCCCAACAGAGGAGGCGGCGGAGTCCGGGGCCGGCG	576
QY	615	CAGTGTGTGTCCAGACGGCGCGGAGAGGGGCCCCAGCGGACCCGACAGCACTGACGA	674
Db	577	CAGTGTGTGTCCAGACGGCGCGGAGAGGGGCCCCAGCGGACCCGACAGCACTGAGYA	635
QY	675	GCCCGCCAGACCTTCTCTGAAAGGCAAAAGTACCTGTGTACATCTCCGGTCCGGGTTCCAGC	734
Db	636	GCCCGCCAGACCTTCTCTGAAAGGCAAAAGTACCTGTGTACATCTCCGGTCCGGGTTCCAGC	695
QY	735	AGACTACTGCAGAGATGGGCGAGGCTTGAGACAGGGCTGTGCATCCCGGCGAGCCCAAGC	794
Db	696	AGACTACTGCAGAGATGGGCGAGGCTTGAGACAGGGCGTGTGCATCCCGGCGG--CCAGGC	754
QY	795	GCTGGCGCGGCACTGCGAGCTGTGAGCTGTGGGCAAGGCCACCGCAGTGTCTTCAAGGAGCT	854
Db	755	GCTGGCGCGGCACTGCGGAGCTGTGGGCAAG--CACCGCAAGTCTCCCTCAAGGAGACT	813
QY	855	GGGCTCTGTGTGTGTGATCAATCAAGTTCTCAAGAGTCTCTATCTGAGAGGCTTCT--GGG	913
Db	814	GGGCTCTGTGTGTGTGATCAATCAAGTTCTCAAGAGTCTCTATCTGAGAGGCTTCTTGGGG	873
QY	914	GCGACTACCTGAGTGGCGGCTGTGTCGAGGCGCTTGGGGGCGGTGTCTCTGACTGAGGCC	973
Db	874	GCGACTACCTGAGTGGCGGCTGTGTCGAGGCGCTTGGGGGCGGTGTCTCTGACTGAGGCC	933
QY	974	TGCGAGAGGCTGTGGCGCGGAGAGGTGTGTCCCTGCTCAGTGTGTGATGAGAGTGA	1033
Db	934	TGCGAGAGGCTGTGTGGCGCGGAGAG--TGTTCGCTTCCGCTCAGTGTGTGATGAGTGA	992
QY	1034	ATGAGGCTGGGCGCGGCGGCTGTGTGCTGATGTAGAGAGAGAGGGGGGCGGCG	1084
Db	993	ATAAGG--TGCGGGGCGCGGCTGTGTGCTGATGTAGAGAGAGAGGGGGGCGGCGSAC	1041
RESULT 10			
BX375652			
LOCUS			
BX375652 Homo sapiens NEUROBLASTOMA cell 25-NORMALIZED Homo sapiens			
DEFINITION			
CDNA clone CS00DC015J24 5-PRIME, mRNA sequence.			
ACCESSION			
BX375652			
VERSION			
BX375652.2 GI:46573220			

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Li, W. B., Gruber, C., Jeesee, J., and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On May 8, 2003 this sequence version replaced gi:30448429.
FEATURES	<p>Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqre@genoscope.cns.fr, web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and clonees into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized.</p> <p>For more information about this cluster, see http://www.genoscope.cns.fr/cdna?cs=CS0DC015DE12QP1&c=9074.r.</p>
SOURCE	<p>Location/Qualifiers</p> <p>1..951</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="CS0DC015YJ24"</p> <p>/library="NEUROBLASTOMA COT 25-NORMALIZED"</p> <p>/csize_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"</p> <p>/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."</p>
ORIGIN	
Query Match	45.1%; Score 848.6; DB 5; Length 951;
Best Local Similarity	98.2%; Pred. No. 4.6e-192;
Matches	908; Conservative 2; Mismatches 8; Indels 7; Gaps 5
Db	49 CGGCCGTCGCCAGAGATCATCGAGCGCAGAAATCTGGCCGGGTTCTGAGCTTGTTCGCC 108
Db	1 CGGCGTCGCGCCAGAGATCATCGAGCGCAGAAATCTGGCGGTTCTGAGCTTGTTCGCC 60
Qy	109 CTCCTCTCCCCGGGAATGCGCTATCCGGGTGACGCCGCCGCCGTCGTGGGAGAGAT 168
Db	61 CTCCTCTCCCCGGGAATGCGCTATCCGGGTGACGCCGCCGCCGTCGTGGGAGAGAT 120
Qy	169 GAGTGCCTGAGCTACTACGGAGATGCTGTTCAGCCGTATGTTGAGATGCTGGCGGG 228
Db	121 GAGTGCCTGAGCTACTACGGAGATGCTGTTCAGCCGTATGTTGAGATGCTGGCGGG 180
Qy	229 CAATGACCGAGTGCAGCTGGAGTCTCTGGCTTTCTGTCTGATGAGGCTCTTGGCGGC 288
Db	181 CAATGACCGAGTGCAGCTGGAGTCTCTGGCTTTCTGTCTGATGAGGCTCTTGGCGGC 240
Qy	289 GCCGAGAGCTTAAGCCCGGAGCCCGGAGCGGCTTAAGCTCTCTGCTGAGAGCGCGC 348
Db	241 GCCGGA-GSTTAAGCCCGGAGCCCGGAGCGGCTTAAGCTCTCTGCTGAGAGCGCGC 299
Qy	349 GGGCAGTGCAGGAGAGCAACTTGGCTGTGGGGCAACTTCTGCGCGTGTGGCGGC 408
Db	300 GNGCAGTGCAGGAGAGCAACTTGGCTGTGGGGCAACTTCTGCGCGTGTGGCGGC 359
Qy	409 CAGCACTGTCTGCGGACCTTGGCGCGCAAGCGCGCGCGGCAAGTCTTCCAAAGCTAT 468
Db	360 CAGCACTGTCTGCGGACCTTGGCGCGCAAGCGCGCGCGGCAAGTCTTCCAAAGCTAT 417
Qy	469 AGCATATGACCTCTCAGCTCTTCAAAGAGACAGAGGTTAGTGCCTGTGCGCGCGAG 528
Db	418 AGCATATGACCTCTCAGCTCTTCAAAGAGACAGAGGTTAGTGCCTGTGCGCGCGAG 477
Qy	529 TCAAGAGTTTCGAATTTCTCAGCAGGGTCAGTGGAGACAGGCTCCCCCCCAACAG 588

|||||
Db 478 TCAGCAGGTTCTGAAATTCACAGAGGTGAGTGGAGAGAGGCTCCCCCAACCAAG 537
|||
Qy 589 CGGAGAGG-6GAGAGTGGGGGCGGCGCAGAGTGTGTCCAGAGCGGGCGGAGAGGGGC 647
|||
Db 538 CGGAGAGGCGAGAGTGGGGCGGCGCAGAGTGTGTCCAGAGCGGGCGGAGAGGGGC 597
|||
Qy 648 CCCAGCGCAGCCCGAGCAGAGTGAAGCCCGCAGACTTCTCTGAAGGCAAGTGAAC 707
|||
Db 598 CCCAGCGCAGCCCGAGCAGAGTGAAGCCCGCAGACTTCTCTGAAGGCAAGTGAAC 657
|||
Qy 708 CTGTGACATCCGCTCCGGGTTGAGAGAGTACTGCGAGCATGGCCAGCTTGAAGCA 767
|||
Db 658 CTGTGACATCCGCTCCGGGTTGAGAGAGTACTGCGAGCATGGCCAGCTTGAAGCA 717
|||
Qy 768 GGGGGTGGCATCCGGGGGCGCCAGGGGCTGGCGGAGCTGAGACGTTTGGGCAAGG 827
|||
Db 718 GGGGGTGGCATCCGGGGGCGCCAGGGGCTGGCGGAGCTGAGACGTTTGGGCAAGG 777
|||
Qy 828 CACCGCAGTGTGCGCTCAAGGAGCTGGGCTGTGTGTGTGACATCAAGTTCTCAGA 887
|||
Db 778 CACCGCAGTGTGCGCTCAAGGAGCTGGGCTGTGTGTGTGACATCAAGTTCTCAGA 836
|||
Qy 888 GCTTCTCTATCTGAAGCGCTTCTGGGGGAGTACTGAGTGGCGCCCTGTCAGAGCCCT 947
|||
Db 837 GCTTCTCTATCTGAAGCGCTTCTGGGGGAGTACTGAGTGGCGCCCTGTCAGAGCC-- 894
|||
Qy 948 GCGGGGGGTGTCTGACTGAGAGCC 972
|||
Db 895 TGGGGGGGTGTCTGACTGAGAGCC 919
|||

RESULT 11
BX342836/c 1166 bp mRNA linear EST 07-APR-2004

LOCUS BX342836 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL006YL09 3-PRIME, mRNA sequence.
ACCESSION BX342836
VERSION BX342836.2 GI:46264629

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1. (bases 1 to 1166)
T.L.M.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 1, 2003 this sequence version replaced gi:30311123.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DL006CF05NP1&c=9074.r.

FEATURES
source 1.1166
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CS0DL006YL09"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 44.4%; Score 836.4; DB 5; Length 1166;
Best Local Similarity 88.8%; Pred. No. 3.9e-189;
Matches 993; Conservative 30; Mismatches 28; Indels 67; Gaps 9;

764 AGCAGGCGTGGCATCCCGGCGGCGCAGAGCGTGGGAGAGCTGAGCTGTTGGGC 823
|||
Db 1139 ACCTTSGAGCGGSKTSCACCGGCGCAGAGTGTGTGACATTTTGGGC 1061
|||
Qy 824 AGGCGCAGAGTGTGCTGCTCAAGGAGCTGGGCTGTGTGTGTGACATCAAGTTCT 883
|||
Db 1080 AGCCCAACGAGKTS-KCGCTCAAGGAGCTGGGCTGTGTGTGTGACATCAAGTTCT 1023
|||
Qy 884 CAGAGCTCTCTATCTGAGAGGCTTCTGGGGGAGCTGAGTGGGCGCTGTCAGAG 943
|||
Db 1022 CAG-GSTCTCTATCTGAGAGGCTTCTGGGGGAGCTGAGTGGGCGCTGTCAGAG 964
|||
Qy 944 CCCTGCGGGGAGTGTCTCTGAGAGGCGCTGAGAG-GGCTGGGCGGAGAGCTGTT 1002
|||
Db 963 -SCTGCGGGGAGTGTCTCTGAGAGGCGCTGAGAGGCTGAGAGGCTGTT 905
|||
Qy 1003 CGCTGCTGTGATGATGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1062
|||
Db 904 CGCTGCTGTGATGATGATGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 845
|||
Qy 1063 ATGAGAGAGG-AGGG 1121
|||
Db 844 ATGAGAGAGGAGAGGG 785
|||
Qy 1122 TGATCCCACTCCCAAGTCTCGGGGCGAGCTTCTCTGGGAGAGAGCATCTCTA----- 1176
|||
Db 784 TGATCCCACTCCCAAGTCTCGGGGCGAGCTTCTCTGGGAGAGAGCATCTCTA----- 725
|||
Qy 1177 -----CCCTT 1182
|||
Db 724 AGAGAGTGTGATCTGAGATCTTGTAGAGCTGGAGAGAGCGGGAGAGAGGCGCTT 665
|||
Qy 1183 GACAGCGCTTCCCAAGAGTGTGGGCTGTGAGGCTTAAAGCTTCCAGCTGAGTTCT 1242
|||
Db 664 GACAGCGCTTCCCAAGAGTGTGGGCTGTGAGGCTTAAAGCTTCCAGCTGAGTTCT 605
|||
Qy 1243 TCCAGAGTCTCTCTACCCCGAGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1302
|||
Db 604 TCCAGAGTCTCTCTACCCCGAGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 545
|||
Qy 1303 TGTATCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1362
|||
Db 544 TGTATCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 485
|||
Qy 1363 TCTGTGATCTTGGGCTGTGTGACAGTGTGACAGAGTGTGACAGAGTGTGACAGT 1422
|||
Db 484 TCTGTGATCTTGGGCTGTGTGACAGTGTGACAGAGTGTGACAGAGTGTGACAGT 425
|||
Qy 1423 CCCCTGTGAGATTTTCCGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 1482
|||
Db 424 CCCCTGTGAGATTTTCCGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 365
|||
Qy 1483 GGCCTAAGGCTTCTCTCTGAGAGCTCTAATTGAGCAGACCAACCTGAGGCTTCA 1542
|||
Db 364 GGCCTAAGGCTTCTCTCTGAGAGCTCTAATTGAGCAGACCAACCTGAGGCTTCA 305
|||
Qy 1543 ACATCACTGAGGAGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 1602
|||
Db 304 ACATCACTGAGGAGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 245
|||
Qy 1603 CAGGCGAGGAGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 1662
|||
Db 244 CAGGCGAGGAGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGG 185
|||
Qy 1663 CTTGCTGTCAAGAGCTGAGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1722
|||

Db 184 CCBCTGCTCCAGAGCTGACATGACTTGACATTGATGTCACCTGAGATGTCACATGG 125
Qy 1723 AGTGTATGGCAGCATCATATCAAGGCTTACTGTGACATGGGCGCAAAACAGTAAC 1782
Db 124 AGTGATATGGCAACATCATCAAGGCTTACT-TTGACATGGGGGSKMAACAGTAAC 66
Qy 1783 AGCCACCTTCTTGAAAGGATGCAAGGCTTGGGG 1820
Db 65 AGCCACCTTCTTGAAAGGATGCAAGGCTTGGGG 28

RESULT 12
BX384908/c 1027 bp mRNA linear EST 27-APR-2004
LOCUS BX384908 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL004YA21 3-PRIME, mRNA sequence.
ACCESSION BX384908
VERSION BX384908.1 GI:30455279
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polyes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DL004YA1INP1&c=9074.r.
FEATURES
source location/Qualifiers
1..1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL004YA21"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 44.0%; Score 829.2; DB 5; Length 1027;
Best Local Similarity 92.9%; Pred. No. 2e-187;
Matches 933; Conservative 3; Mismatches 6; Indels 62; Gaps 4;

Qy 906 CTTCTGGGGGAGTACTAGTGGGCGCCCTGCTGACAGCCCTGGGGGCTTCTCTGAC 965
Db 1001 CTTTGGGGGAGTACTAGTGGGCGCCCTGCTGACAG-SCTGCGGGGCGCTTCTCTGAC 943
Qy 966 TGAGGCCCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGCTGATGATGA 1025
Db 942 TGAGGCCCTGCGAGAGGCTGTGGGCGGAG-GCTGTTCGCTGCTGATGATGA 884
Qy 1026 GGGTGAATATGAGGCTGCGCGCGCGCTGTGTGATGAGAGAGAGGGGGCGCG 1085
Db 883 GGGTGAATATGAGGCTGCGCGCGCGCTGTGTGATGAGAGAGAG-GGGGGGGCGG 825
Qy 1086 CCGGACAGAGGCTCTGATCCAGAGCTGAGAGATTTGATCCACTCCAGTCTCGGG 1145

Db 824 CCGGACAGAGGCTCTGATCCAGAGCTGAGATTTCCACTCCAGTCTCGGG 765
Qy 1146 CCACTTCTCTCTGGAGAGAGCACTCTCTA----- 1176
Db 764 CCACTTCTCTCTGGAGAGAGCACTCTCTA CCCCCAGAGAGAGAGAGAGAGAGAGAG 705
Qy 1177 -----CCCTTGACAGCCCTCCCAAGAGATGTGG 1206
Db 704 TGAGGACTGCGAG 645
Qy 1207 GCTTGAGAGGCTTAAACATTTCCAGTGTGATTTCTTCCAGAGCTCTCTA CCCCCAG 1266
Db 644 GCTTGAGAGGCTTAAACATTTCCAGTGTGATTTCTTCCAGAGCTCTCTA CCCCCAG 585
Qy 1267 TGTGCCCCCTTGAAGCTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
Db 584 TGTGCCCCCTTGAAGCTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
Qy 1327 GCTACAGACTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
Db 524 GCTACAGACTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
Qy 1387 TGTACAGAGTGAACAGCTGCTGATGCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
Db 464 TGTACAGAGTGAACAGCTGCTGATGCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 405
Qy 1447 CTCTGCCCCGAGCTGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1506
Db 404 CTCTGCCCCGAGCTGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
Qy 1507 CCTTAAATTTGACACACACCAACCTGGGCTTACACCACTGAGAGAGAGAGAGAG 1566
Db 344 CCTTAAATTTGACACCAACCTGGGCTTACACCACTGAGAGAGAGAGAGAGAGAG 285
Qy 1567 GTGACATGAGGAGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1626
Db 284 GTGACATGAGGAGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225
Qy 1627 CTTCAATTTTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
Db 224 CTTCAATTTTACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 165
Qy 1687 ACTTGACATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
Db 164 ACTTGACATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 105
Qy 1747 GGGCTACTGTTGACATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1806
Db 104 GGGCTACTGTTGACATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 45
Qy 1807 CAAAGGCTTTGGGGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1850
Db 44 CAAAGGCTTTGGGGGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1

RESULT 13
BM920838 1090 bp mRNA linear EST 12-MAR-2002
LOCUS BM920838
DEFINITION AGENOOURT_6706034 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752092
5', mRNA Sequence.
ACCESSION BM920838
VERSION BM920838.1 GI:19371217
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM12785 row: 1 column: 13
 High quality sequence stop: 625.
 Location/Qualifiers

FEATURES

source

1. 1090
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5752092"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 114"
 /note="Organ: pooled brain, lung, testis; Vector:
 PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 42.8%; Score 805.2; DB 5; Length 1090;
 Best Local Similarity 96.0%; Pred. No. 1.2e-181;
 Matches 881; Conservative 0; Mismatches 28; Indels 9; Gaps 5;

QY 23 CCAGAAAGATCAAAACAGTGGCTGGCGCGCTGGCCAGAGTCACTGAGCGCCAGAT 82
 DB 1 CCAGAAAGATCAAAACAGTGGCTGGCGCGCTGGCCAGAGTCACTGAGCGCCAGAT 60
 QY 83 CTGGCGCGGTTCTAGCTGTTCCGCGCTCCCTCCCGGGAATGGCGCTATCCGGGTGA 142
 DB 61 CTGGCGCGGTTCTAGCTGTTCCGCGCTCCCTCCCGGGAATGGCGCTATCCGGGTGA 120
 QY 143 CCCCCGCGCGCTGGAGAGAGATGAGTCTGACTACTAGCGGATGCTGTGCTTC 202
 DB 121 CCCCCGCGCGCTGGAGAGAGATGAGTCTGACTACTAGCGGATGCTGTGCTTC 180
 QY 203 ACCGTATGTTGAGGTGGTGGCGGCACTGACCGAGTGGAGCTCTCTGCGCT 262
 DB 181 ACCGTATGTTGAGGTGGTGGCGGCACTGACCGAGTGGAGCTCTCTGCGCT 240
 QY 263 TTCTGCTGAGAGAGCTCTGGCGCGCGGAGCTTACCGCGCGCGGAGCGGCTAG 322
 DB 241 TTCTGCTGAGAGAGCTCTGGCGCGCGGAGCTTACCGCGCGCGGAGCGGCTAG 300
 QY 323 AGCTCTGCTGAGAGAGCGCGCGGAGAGTGGCGGAGCAACTCTGCGGCTGAG 382
 DB 301 AGCTCTGCTGAGAGAGCGCGCGGAGAGTGGCGGAGCAACTCTGCGGCTGAG 360
 QY 383 GGCAACTCTGCGCGCTGGCGCGCGCACTGCTGCTGCGGCGGCAAGCGGC 442
 DB 361 GGCAACTCTGCGCGCTGGCGCGCGCACTGCTGCTGCGGCGGCAAGCGGC 420
 QY 443 GCGCGCGAGTCTCCGAGAGAGCTATAGCTATGGCACTCCAGCTCTTCAAGAGAGAG 502
 DB 421 GCGCGCGAGTCTCCGAGAGAGCTATAGCTATGGCACTCCAGCTCTTCAAGAGAGAG 480
 QY 503 AGGATAGCTGCGCGCTGGCGGAGTCAAGCAAGTCTGCAAAATCTCAGAGGGTCACT 562
 DB 481 AGGATAGCTGCGCGCTGGCGGAGTCAAGCAAGTCTGCAAAATCTCAGAGGGTCACT 540
 QY 563 GGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGTCTGGGCGCGGCTCAGTGTG 622
 DB 541 GGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGTCTGGGCGCGGCTCAGTGTG 600

QY 623 GTCCAGACCGCGCGGAGAGAGGCGGCCACGCCGACCCGACAGAGTCAAGCCGCCCA 682
 DB 601 GTCCAGACCGCGCGGAGAGAGGCGGCCACGCCGACCCGACAGAGTCAAGCCGCCCA 660
 QY 683 GACCTCTCTTGAAGGCAAGTGAAGTGAATCTGAGTCACTCCGCGCTTCGAGCAGATGACT 742
 DB 661 GACCTCTCTTGAAGGCAAGTGAAGTGAATCTGAGTCACTCCGCGCTTCGAGCAGATGACT 720
 QY 743 GCGAGCATGGGCGCAGCTTGGAGCAGGCGGTGGCATCCGCGCGGCCGACGCGTGGCGC 802
 DB 721 GCGAGCATGGGCGCAGCTTGGAGCAGGCGGTGGCATCCGCGCGGCCGACGCGTGGCGC 780
 QY 803 GCGAGCTGACGCTGTTT--GGGACGCGCAGCGAGTGTCT--CGCTCAAGGAGACTGGGCGC- 858
 DB 781 GCGAGCTGACGCTGTTTGGGACGCGCAGCGAGTGTCTCCGCTCAAGGAGACTGGGCGC 840
 QY 859 TCTGTGTTTGTGACAT--CAAGTCTCAGAG--CTCTCTATCTGAGCGCTTCTGCGG 913
 DB 841 TCTGTGTTTGTGACATCAAGTCTCAGAGGCTCTCCCTATCTGAGCGCTTCTGCGG 900
 QY 914 GCGACTACCTGAGTGGCG 931
 DB 901 GGGCGACCTACCTGAGG 918

RESULT 14

BM925969

LOCUS

DEFINITION

BM925969 1089 bp mRNA linear EST 12-MAR-2002
 AGENCOURT_6649774 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764272
 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM12817 row: h column: 01
 High quality sequence start: 14
 High quality sequence stop: 586.
 Location/Qualifiers

FEATURES

source

1. 1089
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5764272"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 114"
 /note="Organ: brain; Vector: PCMV-SPORT6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 41.6%; Score 782.8; DB 5; Length 1089;
Best Local Similarity 95.9%; Pred. No. 2,7e-176;
Matches 836; Conservative 0; Mismatches 32; Indels 4; Gaps 3;

QY CATCGACGCGCAATCTGGCGGGGTTCTGAGCTTGTCGCTCCCTCCCGGGAAAG 126
DB CATCGACGCGCAATCTGGCGGGGTTCTGAGCTTGTCGCTCCCTCCCGGGAAAG 73
QY GCGCTATCCGGGTGACACCCCGGCGCGTGTGGAGAGAGATGAGCTGACTATAC 186
DB GCGCTATCCGGGTGACACCCCGGCGCGTGTGGAGAGAGATGAGCTGACTATAC 133
QY GGGATGCTGCTGCTTACCTATGTTCTGAGTGTGGCGGGCACTGACCTGAGCGAG 246
DB GGGATGCTGCTGCTTACCTATGTTCTGAGTGTGGCGGGCACTGACCTGAGCGAG 193
QY CTGAGCTCTGAGCTTTCTGCTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCCCG 306
DB CTGAGCTCTGAGCTTTCTGCTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCCCG 253
QY GCGCGAGCGGCTTAGAGCTCTGCTGAGAGCTGAGCGCGCGGAGGCTGAGCGAGAG 366
DB GCGCGAGCGGCTTAGAGCTCTGCTGAGAGCTGAGCGCGCGGAGGCTGAGCGAGAG 313
QY AACCTGCGGCTGCTGGGGCAATCTCTGCGGTGCTGGCGCGCACTGCTGCGGAC 426
DB AACCTGCGGCTGCTGGGGCAATCTCTGCGGTGCTGGCGCGCACTGCTGCGGAC 373
QY CTGCGCGCAAGCGCGCGCGGCGGAGTGTCTCCAGAAAGCTATAGTATGACCTCCAG 486
DB CTGCGCGCAAGCGCGCGCGGCGGAGTGTCTCCAGAAAGCTATAGTATGACCTCCAG 433
QY TCTTCAAGAGAGAGAGAGGAGTGTGCGGCTGCGGAGTCAAGCACTTCTGCAAT 546
DB TCTTCAAGAGAGAGAGAGGAGTGTGCGGCTGCGGAGTCAAGCACTTCTGCAAT 493
QY TCTCAGAGAGGCTAGTGGAGAGAGAGTCTCCCGGCAACAGAGGCGGCGGAGTCCG 606
DB TCTCAGAGAGGCTAGTGGAGAGAGAGTCTCCCGGCAACAGAGGCGGCGGAGTCCG 553
QY GCGCGCGCGAGTGTGGTGCAGAGCGGCGGAGAGAGGCGCGGAGCGGCGGAGCGAG 666
DB GCGCGCGCGAGTGTGGTGCAGAGCGGCGGAGAGAGGCGCGGAGCGGCGGAGCGAG 613
QY CAGTCAAGAGCGCGGAGAGCTTCTCTGAGAGGCAAGTGAATGTGATCTCGGCTCCG 673
DB CAGTCAAGAGCGCGGAGAGCTTCTCTGAGAGGCAAGTGAATGTGATCTCGGCTCCG 673
QY GTTCGAGCAGAGTATGCGAGAGTGGGCGGAGCTTGGAGAGAGGCGGTGGATCCCGGG 786
DB GTTCGAGCAGAGTATGCGAGAGTGGGCGGAGCTTGGAGAGAGGCGGTGGATCCCGGG 733
QY CCGCAGGCGCTGAGCGGCGGAGTGAAG-1GTTTGGGAGAGCGGAGGAGTGTGAGCT 844
DB CCGCAGGCGCTGAGCGGCGGAGTGAAG-1GTTTGGGAGAGCGGAGGAGTGTGAGCT 793
QY CAAAGGAGCTGAGCT--CTGTGTGTTTGTGACATCAAGTTCTGAGAGTCTCTATCTGA 902
DB CAAAGGAGCTGAGCT--CTGTGTGTTTGTGACATCAAGTTCTGAGAGTCTCTATCTGA 853
QY CGGCTTCTGGGGGAGTACTGAGTGGCGCC 934
DB CGGCTTCTGGGGGAGTACTGAGTGGCGCC 885

RESULT 15
LOCUS BX381774/c 993 bp mRNA linear EST 28-APR-2004
DEFINITION BX381774 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1069YH17 3-PRIME, mRNA sequence.
ACCESSION BX381774
VERSION BX381774.2 GI:46831712
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 993)
TITLE Li, W.B., Gruber, C., Jessee, J., and Polajnar, D.
JOURNAL Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30456994.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1069YH17&c=9074.r.

FEATURES
Location/Qualifiers
1..993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1069YH17"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 41.2%; Score 775; DB 5; Length 993;
Best Local Similarity 91.9%; Pred. No. 1.9e-174;
Matches 896; Conservative 9; Mismatches 6; Indels 64; Gaps 6;

QY TCTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGCTGCTGCTGCTGCTG 1018
DB TCTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGCTGCTGCTGCTGCTG 914
QY TGGATAGGCTGCTATGAGGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1078
DB TGGATAGGCTGCTATGAGGCTGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 856
QY GCGCGCGCGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1138
DB GCGCGCGCGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 796
QY CTCGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1176
DB CTCGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 736
QY CTCGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1199
DB CTCGCGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
QY GATGTGGGCTGAGGCGCTAAACATTTCCAGTGAATTTCTTCCAGACTCTCTCTAC 1259
DB GATGTGGGCTGAGGCGCTAAACATTTCCAGTGAATTTCTTCCAGACTCTCTCTAC 616
QY CCGCAGGTGCGCCCTTAGCTCCGAGGCGGCGGCTGGCTGTATCTCAGAAAGGAG 1319
DB CCGCAGGTGCGCCCTTAGCTCCGAGGCGGCGGCTGGCTGTATCTCAGAAAGGAG 556
QY GGGCAGAGTAACTCACTCAAAAGGCGCGCTGCACTTTGATCTTCTGAGCTTG 1379
DB GGGCAGAGTAACTCACTCAAAAGGCGCGCTGCACTTTGATCTTCTGAGCTTG 496
QY TCTGCACTGTCAAGTGAACACTGCTGATGCTCAACATGCGCCCTCTGAGATCTTC 1439

```
Db      495 TCTGCACTGTCAAGGTGACACACATGCTCATGCTCACTGCCCCGCTGAGATCTTC 436
QY      1440 CTTGGGCTCTGCGCTTGCCCTTGCTCCAGACACACTTCTTGCCCTTAAGGCTTCTCT 1499
Db      435 CTTGGGCTCTGCGCTTGCCCTTGCTCCAGACACACTTCTTGCCCTTAAGGCTTCTCT 376
QY      1500 CTCAGGACCTCTAATTGTTGACCAACCACTGCGCTTCAGCCACATCAGTGGGCACTGG 1559
Db      375 CTCAGGACCTCTAATTGTTGACCAACCACTGCGCTTCAGCCACATCAGTGGGCACTGG 316
QY      1560 AGCTGGGGTGCACATGGGGGCTGTGCTACCTTGCCACACATCTTCAGCCAGCCAGGGCCC 1619
Db      315 AGCTGGGGTGCACATGGGGGCTGTGCTACCTTGCCACACATCTTCAGCCAGCCAGGGCCC 256
QY      1620 TGCCAGCTTCAATTTCAGACCTGACTCTCTCACTTCCCTGCTGTGCCAGAGCTG 1679
Db      255 TGCCAGCTTCAATTTCAGACCTGACTCTCTCACTTCCCTGCTGTGCCAGAGCTG 196
QY      1680 AACATAGACTTGCACCTTGGATGTTCACCTGGAGTGTCAATGGAGTGTATGGCAGCATC 1739
Db      195 AACATAGACTTGCACCTTGGATGTTCACCTGGAGTGTCAATGGAGTGTATGGCAGCATC 136
QY      1740 ATACCAAGGCTTACTGTGCAATGGGGCCAAAACGATTAACAGCCACTTCTTGAAA 1799
Db      135 ATACCAAGGCTTACTGTGCAATGGGGCCAAAACGATTAACAGCCACTTCTTGAAA 76
QY      1800 GGGATGCAAAAGGCTTGGGGGTGATGGAAGAAGCTTTTA-CAATGATACCAATTAAA 1858
Db      75 GGGATGCAAAAGGCTTGGGGGTGATGGAAGAAGCTTTTANCAATGANTNCCAAITTAIV 16
QY      1859 CTGCCCCTGAAAGGG 1873
Db      15 VNVCCGNGGMMAGGG 1
```

Search completed: March 23, 2005, 16:09:39
Job time : 6757.64 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OW nucleic - nucleic search, using SW model

Run on: March 23, 2005, 07:12:12 ; Search time 8476.2 Seconds
(without alignments)
10764.399 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 1883
Sequence: 1 aggcgcataacatagagaag.....ctgaaagggcatagctg995 1883

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:.*
1: gb_ba:.*
2: gb_hgt:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_str:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1883	100.0	1883	6	BD093312 Apoptosis
2	1883	100.0	1883	6	AK075328 Homo sapi
3	1754.2	93.2	1951	9	BC027930 Homo sapi
4	1722.2	91.5	2012	9	AY125488 Homo sapi
5	1701.4	90.4	2005	9	BC013372 Homo sapi
6	1681.4	89.3	1924	6	AX431308 Sequence
7	1566.8	66.7	1659	9	AK130203 Homo sapi
8	1180.8	62.7	1230	6	AX322754 Sequence
9	1103	58.6	95663	9	AC010247 Homo sapi
10	979.4	52.0	981	6	CO719642 Sequence
11	979.4	52.0	981	9	AF443591 Homo sapi
12	971.4	51.6	981	9	AF457575 Homo sapi
13	946.4	50.3	1106	6	AX364852 Sequence
14	909	48.3	909	6	BD093311 Apoptosis
15	902.4	47.9	1084	6	AX364851 Sequence
16	763.4	40.5	993	10	AF457576 Mus muscu
17	763.4	40.5	993	10	AF543541 Mus muscu
18	719	38.2	847	6	AK379637 Sequence
19	628	33.4	1067	6	AK339362 Sequence

20	615	32.7	626	6	BD149917 Primer fo
21	615	32.7	626	6	AX869855 Sequence
22	457.2	24.3	603	6	AX136449 Sequence
23	457.2	24.3	603	6	BD123689 Secretary
24	434.2	23.1	1540	10	BC037043 Mus muscu
25	388	20.6	190659	2	AC120393 Mus muscu
26	388	20.6	226909	9	AC079490 Mus muscu
27	365.6	19.4	111084	9	AC006486 Homo sapi
28	355.8	18.9	381	6	BD154885 Primer fo
29	355.8	18.9	381	6	AX874823 Sequence
30	301.4	16.0	303	6	AX431298 Sequence
31	296.8	15.8	441	6	CQ460291 Sequence
32	254	13.5	167108	2	AC068283 human STS A
33	221.6	11.8	222	11	G20341 human STS S
34	213	11.3	250	11	G31114 human STS S
35	208	11.0	2624	5	BC084232 Xenopus l
36	207.2	11.0	143720	2	AC145501 Canis fam
37	206.2	11.0	1139	6	BD227199 Protein f
38	206.2	11.0	1139	6	AX008886 Sequence
39	206.2	11.0	1142	6	AR477511 Sequence
40	206.2	11.0	1830	6	BD157757 Primer fo
41	206.2	11.0	1830	6	AX879289 Sequence
42	206.2	11.0	1830	9	AK022531 Homo sapi
43	206.2	11.0	2261	6	BD156359 Primer fo
44	206.2	11.0	2261	6	AX397469 Sequence
45	206.2	11.0	2261	6	AX876867 Sequence

ALIGNMENTS

RESULT 1	BD093312	1883 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093312				
DEFINITION	Apoptosis Relating Factor.				
ACCESSION	BD093312.1	GI:22638900			
VERSION	BD093312.1	GI:22638900			
KEYWORDS	WO 0104300-A/2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Okazaki, T., Isogai, T., Nishikawa, T., Kawai, Y., Miyoshi, S. and Sato, S. (1999) J. Biol. Chem. 274:15956-15966				
AUTHORS	Okazaki, T., Isogai, T., Nishikawa, T., Kawai, Y., Miyoshi, S. and Sato, S.				
TITLE	Apoptosis Relating Factor				
JOURNAL	Patent: WO 0104300-A/2 18-JAN-2001; HELIX RESEARCH INSTITUTE, TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI, SUSUMU SATO				
COMMENT	OS Homo sapiens (human) PN WO 0104300-A/2 PR 06-JUL-2000 WO 2000JP004516 PR 08-JUL-1999 JP 99P 194179, 18-OCT-1999 US 60/159566 TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI, PI SUSUMU SATO PC C12N5/12, C07K14/47, C12N5/10, C12N1/21, C12N1/19, C12N1/15, C12P21/ PC 02, C07K16/18, C12P21/08, G01N33/53, G01N33/577 PC CC Key location/Qualifiers FH Key location/Qualifiers FT CDS (124) (1101).				
FEATURES	Source	location/Qualifiers			
ORIGIN	1..1883 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"				
Query Match	100.0%; Score 1883; DB 6; Length 1883;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1883; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				

QY 1 AGCGTAATATAGAGAGAGTGCAGAAAATTCAAAACAGAGTGCCTGCGCCGCTGCCC 60
 Db 1 AGCGTAATATATAGAGAGAGTGCAGAAAATTCAAAACAGAGTGCCTGCGCCGCTGCCC 60
 QY 61 AGAGATCATCGAGCCGCAAAATTCGCGCCGCTGCTGAGCTTGTTCGCGCTCCCTCCCGG 120
 Db 61 AGAGATCATCGAGCCGCAAAATTCGCGCCGCTGCTGAGCTTGTTCGCGCTCCCTCCCGG 120
 QY 121 GGAATGCGCTATCCGGGTGAGACCCCGGCCCCGCTGCGAGAGAGATGATGCTGAGAC 180
 Db 121 GGAATGCGCTATCCGGGTGAGACCCCGGCCCCGCTGCGAGAGAGATGATGCTGAGAC 180
 QY 181 TACTAGGGATGCTGCTTCAACGTAATGTTGAGAGTGTGGGCGGGCAACTGACCGAG 240
 Db 181 TACTAGGGATGCTGCTTCAACGTAATGTTGAGAGTGTGGGCGGGCAACTGACCGAG 240
 QY 241 TCGGAGCTGAGAGCTCTGGGCTTTCTGCTGAGATGAGAGCTTCGCGCCCGGAGGCTTA 300
 Db 241 TCGGAGCTGAGAGCTCTGGGCTTTCTGCTGAGATGAGAGCTTCGCGCCCGGAGGCTTA 300
 QY 301 GCCCGGCGCCGAGCGGCTGAGAGCTCTGCTGAGAGCTGAGAGCGCGCGAGTGCAGC 360
 Db 301 GCCCGGCGCCGAGCGGCTGAGAGCTCTGCTGAGAGCTGAGAGCGCGCGAGTGCAGC 360
 QY 361 GAGAGCAACTGCGAGCTGCTGGGGCAACTCTGCGGGTCTGCGCCGAGCACTGCTG 420
 Db 361 GAGAGCAACTGCGAGCTGCTGGGGCAACTCTGCGGGTCTGCGCCGAGCACTGCTG 420
 QY 421 CCGCACTGCGCGGAGAGCGGCGCCGCGCAAGTCTTCCAGAAAGCTATGCTATGCGAC 480
 Db 421 CCGCACTGCGCGGAGAGCGGCGCCGCGCAAGTCTTCCAGAAAGCTATGCTATGCGAC 480
 QY 481 TCCAGCTCTTCAAAAGAGAGACAGAGGAGTCTGCTGCGCGCTGCGAGTCAAGCACTTCT 540
 Db 481 TCCAGCTCTTCAAAAGAGAGACAGAGGAGTCTGCTGCGCGCTGCGAGTCAAGCACTTCT 540
 QY 541 GCAAAATTTCTAGCAGAGGCTGAGAGAGAGAGCTGCTCCCGCAACAGCGCGAGCGGCG 600
 Db 541 GCAAAATTTCTAGCAGAGGCTGAGAGAGAGAGCTGCTCCCGCAACAGCGCGAGCGGCG 600
 QY 601 AGTCGGGGCGCGCGCGCTGCTGCTGAGAGAGAGCGCGCGAGAGAGCGCGCGCGCG 660
 Db 601 AGTCGGGGCGCGCGCGCTGCTGCTGAGAGAGAGCGCGCGAGAGAGCGCGCGCGCG 660
 QY 661 CAGCAGCACTGAGAGCGCGCGCAGAGCTTCTTGAAGCAAAAGTGAACCTGTGACATCCG 720
 Db 661 CAGCAGCACTGAGAGCGCGCGCAGAGCTTCTTGAAGCAAAAGTGAACCTGTGACATCCG 720
 QY 721 CTCGGGGTTTGAAGCAGAGTACTGCGAGCATGAGGCGCAGCTTGGAGCAGGCGCTG 780
 Db 721 CTCGGGGTTTGAAGCAGAGTACTGCGAGCATGAGGCGCAGCTTGGAGCAGGCGCTG 780
 QY 781 CCGCGCGCGCGCGCGCTGCTGCTGAGAGAGTGTGTTGGAGAGCGCGCAGAGTGTG 840
 Db 781 CCGCGCGCGCGCGCGCTGCTGCTGAGAGAGTGTGTTGGAGAGCGCGCAGAGTGTG 840
 QY 841 CGCTCAAGGAGACTTGGGCTCTGCTGTTGATCATCAAGTTCAGAGCTCTCTATCTG 900
 Db 841 CGCTCAAGGAGACTTGGGCTCTGCTGTTGATCATCAAGTTCAGAGCTCTCTATCTG 900
 QY 901 GACGCTTCTGGGGCAGACTTACTGAGTGTGCGCTTCTGAGAGCGCTTGGCGGGCGTGT 960
 Db 901 GACGCTTCTGGGGCAGACTTACTGAGTGTGCGCTTCTGAGAGCGCTTGGCGGGCGTGT 960
 QY 961 CTGACTGAGAGCGCTTGGAGAGGCTGTGGGCGCGGAGAGGCTTGTGCTGTGCTGCTG 1020
 Db 961 CTGACTGAGAGCGCTTGGAGAGGCTGTGGGCGCGGAGAGGCTTGTGCTGTGCTGCTG 1020
 QY 1021 GATGAGGCTGACTATGAGGCTGCGCGCGCGCTGCTGCTGCTGATGAGAGAGGAGG 1080
 Db 1021 GATGAGGCTGACTATGAGGCTGCGCGCGCGCTGCTGCTGCTGATGAGAGAGGAGG 1080

QY 1081 CGCGCGCGCAGAGAGGCTCTGATTCAGAGATCGGAGAGATTCAGCTCCCAAGTCT 1140
 Db 1081 CGCGCGCGCAGAGAGGCTCTGATTCAGAGATCGGAGAGATTCAGCTCCCAAGTCT 1140
 QY 1141 CGGGGCAACCTTCTCTGGGAGAGAGCAGCATCTTACCCTTGAAGCGCCCTCCACAG 1200
 Db 1141 CGGGGCAACCTTCTCTGGGAGAGAGCAGCATCTTACCCTTGAAGCGCCCTCCACAG 1200
 QY 1201 ATGTGGGCTCTGAGGCTTAAACATATTCAGCTGATTTCTTCCAGACTCTCTTACC 1260
 Db 1201 ATGTGGGCTCTGAGGCTTAAACATATTCAGCTGATTTCTTCCAGACTCTCTTACC 1260
 QY 1261 CCCAGGTGTGCCCCCTTGAAGCTTCCGAGAGCGGGGCTGGGCTGTATCTCAGAGAGAG 1320
 Db 1261 CCCAGGTGTGCCCCCTTGAAGCTTCCGAGAGCGGGGCTGGGCTGTATCTCAGAGAGAG 1320
 QY 1321 GGCACAGCTACACATCAACCAAGGCGCCGCTGCACTTGTATCTGTGATCTTGGGCTGT 1380
 Db 1321 GGCACAGCTACACATCAACCAAGGCGCCGCTGCACTTGTATCTGTGATCTTGGGCTGT 1380
 QY 1381 CTGCACTGTCAAGGTGACACACATCTGCTCATGCTCACACTGCCCTGCTGAGATCTTCC 1440
 Db 1381 CTGCACTGTCAAGGTGACACACATCTGCTCATGCTCACACTGCCCTGCTGAGATCTTCC 1440
 QY 1441 CTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 Db 1441 CTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
 QY 1501 TCGAGGACTCTAATTTGACACCAACCAACCTGCGGCTTCCAGCACAATCACTGAGCACTGGA 1560
 Db 1501 TCGAGGACTCTAATTTGACACCAACCAACCTGCGGCTTCCAGCACAATCACTGAGCACTGGA 1560
 QY 1561 GCTGGGGTGCACATGAGGCGCTGCTCACTTGGCCACACATCTCCAGCCAGCGAGGCGCT 1620
 Db 1561 GCTGGGGTGCACATGAGGCGCTGCTCACTTGGCCACACATCTCCAGCCAGCGAGGCGCT 1620
 QY 1621 GCCCAGCTTCAATTTACAGACCTGACTCTCTCACTTCCCTGCTGCTGCTGCTGCTG 1680
 Db 1621 GCCCAGCTTCAATTTACAGACCTGACTCTCTCACTTCCCTGCTGCTGCTGCTGCTG 1680
 QY 1681 ACATGACTTGCACCTGATGATGACCTGAGAGTGTCAATGAGAGTGTATGAGCAGATCA 1740
 Db 1681 ACATGACTTGCACCTGATGATGACCTGAGAGTGTCAATGAGAGTGTATGAGCAGATCA 1740
 QY 1741 TACCAAGGCTTACTGTTGACATGAGGCGCAAAACAGTAACAGCCACTTCTTGAAG 1800
 Db 1741 TACCAAGGCTTACTGTTGACATGAGGCGCAAAACAGTAACAGCCACTTCTTGAAG 1800
 QY 1801 GGAATGCAAAAGCTTGGGGGTGATGAGAAAGACCTTTTACAAATGATACCAATTAACT 1860
 Db 1801 GGAATGCAAAAGCTTGGGGGTGATGAGAAAGACCTTTTACAAATGATACCAATTAACT 1860
 QY 1861 GCCCTGGAAGAGGCAATAGGTGGG 1883
 Db 1861 GCCCTGGAAGAGGCAATAGGTGGG 1883

RESULT 2
 AK075328
 LOCUS 1883 bp mRNA linear PRI 03-SEP-2002
 DEFINITION Homo sapiens cDNA P58C0004 F15, clone NT7RM1000558, weakly similar to Homo sapiens death effector domain-containing testicular molecule mRNA.
 ACCESSION AK075328
 VERSION AK075328.1 GI:22761346
 KEYWORDS oligo capping, f15 (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 Ota, T., Nishikawa, T., Suzuki, Y., Kawai-Hiro, Y., Hayashi, K., Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T.,

TITLE Nakamura, Y., Nagahari, K., Sugano, S. and Iwagai, T.
JOURNAL HRI human cDNA sequencing project
REFERENCE 2 (bases 1 to 1883)
AUTHORS Iwagai, T. and Yamamoto, J.
TITLES Direct Submission
JOURNAL Submitted (20-MAR-2002) Takao Iwagai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass
sequencing, clone selection and full insert sequencing; Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction; Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
Location/Qualifiers
1. 1883

COMMENT

FEATURES

Source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM1000558"
/cell_type="NT2"
/cell_line="teratocarcinoma"
/clone_id="NT2RM1"
/note="Cloning vector: pUC19FL3-mRNA from undifferentiated NT2
neutrophil precursor cells"
124. 1104
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11551.1"
/db_xref="GI:22761347"
/translation="MALSGSTPAFCMEDECLDYGMLSLHMEFVVGQLTTECELT
LAFLEAPGAGAGLARGSLLELEERGGESNLGOLLEVLARHDLPLH
ARRRRVSPERSYGTSSSKRTESCRBRROSSANSOOWEMGSPTRORS
KRRPSGARRRRGAPAPAPQOQSPAPSPSEKGTICIRLRVAPYEHHPALBQGYA
SRPQALROLDFGATAVRSRDLDSVVCIDIFSELSTLDFAWSYLSGALLQALR
GVFLTEALREAVGHEAVRLVSYDEADYEAGRRRLIMEEGRRPTEAS"

CDS

Query Match 100.0%; Score 1883; DB 9; Length 1883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 1 AGGCGTAAATATAGAGAGGTGCGAAGAAAGATCCAAAACAAGTGGCTGCGGCGCTCGCCC 60
DB 1 AGGCGTAAATATAGAGAGGTGCGAAGAAAGATCCAAAACAAGTGGCTGCGGCGCTCGCCC 60
QY 61 AGGAGTCATGAGAGCGCAAAATCTGGCCGGGCTTCTGAGCTTGTCCGCTTCCTCCCG 120
DB 61 AGGAGTCATGAGAGCGCAAAATCTGGCCGGGCTTCTGAGCTTGTCCGCTTCCTCCCG 120
QY 121 GGAATGGCGCTATCCGGGTGAGACCCCGGCGGCTGGAGAGAGATGAGTGGCTGGAC 180
DB 121 GGAATGGCGCTATCCGGGTGAGACCCCGGCGGCTGGAGAGAGATGAGTGGCTGGAC 180
QY 181 TACTACGGAGATGCTGCTTCAACGATATGTTGAGAGTGGTGGGCAACTGACCGAG 240
DB 181 TACTACGGAGATGCTGCTTCAACGATATGTTGAGAGTGGTGGGCAACTGACCGAG 240
QY 241 TGGAGCTGAGAGCTCTGCTCTTCTGCTGATGAGGCTCTGAGCGCCGCGAGGCTTA 300
DB 241 TGGAGCTGAGAGCTCTGCTCTTCTGCTGATGAGGCTCTGAGCGCCGCGAGGCTTA 300
QY 301 GCGCGGGCGCGACGGGCTGAGAGCTCTGCTGAGAGTGGAGCGCCGCGGCAATGGGCG 360
DB 301 GCGCGGGCGCGACGGGCTGAGAGCTCTGCTGAGAGTGGAGCGCCGCGGCAATGGGCG 360
QY 361 GAGAGCAACTGCTGGCTGCTGGGCAACTCTGCGGCTGGGCGCCGCGCAAGACTCTG 420
DB 361 GAGAGCAACTGCTGGCTGCTGGGCAACTCTGCGGCTGGGCGCCGCGCAAGACTCTG 420
QY 421 CCGGCACTGGCGCGCAAGCGGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGGAC 480
DB 421 CCGGCACTGGCGCGCAAGCGGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGGAC 480

QY 481 TCCAGCTCTTCAAGAGAGACAGAGGTAGCTGCGCGCTGCGCGGAGTCAACAGTTCT 540
DB 481 TCCAGCTCTTCAAGAGAGACAGAGGTAGCTGCGCGCTGCGCGGAGTCAACAGTTCT 540
QY 541 GCAAAATTTTCAGAGAGGTCTGATGGAGACAGGCTCCCCCAACCAAGCGGCGGCG 600
DB 541 GCAAAATTTTCAGAGAGGTCTGATGGAGACAGGCTCCCCCAACCAAGCGGCGGCG 600
QY 601 AGTCGGGCGCGGCGCAGTGGTGGTGCAGAGCGGCGGAGAGAGGCGCCAGCGCACCC 660
DB 601 AGTCGGGCGCGGCGCAGTGGTGGTGCAGAGCGGCGGAGAGAGGCGCCAGCGCACCC 660
QY 661 CACAGCAGTCAAGAGCGCGCAGACCTTCTCTGAAAGCAAAATGACCTTGAATCCG 720
DB 661 CACAGCAGTCAAGAGCGCGCAGACCTTCTCTGAAAGCAAAATGACCTTGAATCCG 720
QY 721 CTCGGGTTTCAGAGAGTATCTGAGCATGGGCGCAGCTTGGAGCAGGCGTGGCATCC 780
DB 721 CTCGGGTTTCAGAGAGTATCTGAGCATGGGCGCAGCTTGGAGCAGGCGTGGCATCC 780
QY 781 CCGCGGCGCGGCGCTGCGGCGGCGCAGCTGAGAGTGTGGGAGGCGCACCGAGTCTG 840
DB 781 CCGCGGCGCGGCGCTGCGGCGGCGCAGCTGAGAGTGTGGGAGGCGCACCGAGTCTG 840
QY 841 CGCTCAAGGAGACTGGGCTCTGTGTTTGTGATCATCAAGTTCTCAGAGCTCTCTATCTG 900
DB 841 CGCTCAAGGAGACTGGGCTCTGTGTTTGTGATCATCAAGTTCTCAGAGCTCTCTATCTG 900
QY 901 GAGCGCTTCTGGGCGGAGCTTCTGAGTGGCGCTCTGCTGAGGCGCTGCGGCGGCTGTC 960
DB 901 GAGCGCTTCTGGGCGGAGCTTCTGAGTGGCGCTCTGCTGAGGCGCTGCGGCGGCTGTC 960
QY 961 CTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGAGGCTTGTGCTGCTGTGCTGAGT 1020
DB 961 CTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGAGGCTTGTGCTGCTGTGCTGAGT 1020
QY 1021 GATGAGGCTGACTATGAGAGCTGAGCGGCGCGGCTGTTGTGATGAGAGAGAGGAGG 1080
DB 1021 GATGAGGCTGACTATGAGAGCTGAGCGGCGCGGCTGTTGTGATGAGAGAGAGGAGG 1080
QY 1081 CCGCGCGCGGAGAGAGGCTCTGATCCAGAGCTGGAGAGTGTATCCCACTCAAGTCT 1140
DB 1081 CCGCGCGCGGAGAGAGGCTCTGATCCAGAGCTGGAGAGTGTATCCCACTCAAGTCT 1140
QY 1141 CCGGCGCACTTCTCTGCGGAGAGAGCAATCTTCAACCTTGAACAGCCCTCCACAGG 1200
DB 1141 CCGGCGCACTTCTCTGCGGAGAGAGCAATCTTCAACCTTGAACAGCCCTCCACAGG 1200
QY 1201 ATGTGGGCTCTGAGGCTTAACCAATTTCCAGCTGAGTGTCTTCCAGACTCTCTCAAC 1260
DB 1201 ATGTGGGCTCTGAGGCTTAACCAATTTCCAGCTGAGTGTCTTCCAGACTCTCTCAAC 1260
QY 1261 CCGAGGTGTCGCCCTTGAAGCTTCGAGAGCGGAGGCTGGGCTGTATCTCAGAAAGGAG 1320
DB 1261 CCGAGGTGTCGCCCTTGAAGCTTCGAGAGCGGAGGCTGGGCTGTATCTCAGAAAGGAG 1320
QY 1321 GGCACAGCTACACATCAACCAAGGCGCCCTGCAATTTATCTGATCTTGGGCTGT 1380
DB 1321 GGCACAGCTACACATCAACCAAGGCGCCCTGCAATTTATCTGATCTTGGGCTGT 1380
QY 1381 CTGCACTGTACAGAGTGCACACACTGCTATCTCACTGCTGCTGCTGCTGCTGCTGCT 1440
DB 1381 CTGCACTGTACAGAGTGCACACACTGCTATCTCACTGCTGCTGCTGCTGCTGCTGCT 1440
QY 1441 CTGGGCTCTGCGCTGCGCTGCTTCCAGAGCACTTCTTGGGCTTAAGGAGCTTCTCTC 1500
DB 1441 CTGGGCTCTGCGCTGCGCTGCTTCCAGAGCACTTCTTGGGCTTAAGGAGCTTCTCTC 1500
QY 1501 TCAAGACTTATATTTGACCAACCAACCACTGGGCTTCAAGCAATCAAGTGGGCACTGGA 1560
DB 1501 TCAAGACTTATATTTGACCAACCAACCACTGGGCTTCAAGCAATCAAGTGGGCACTGGA 1560

```

QY 1561 GCTGGGGGTGACATGGGGGCTGCTCACTTGGCCCAACATCTCCAGCCAGCCGGCCCT 1620
DB 1561 GCTGGGGGTGACATGGGGGCTGCTCACTTGGCCCAACATCTCCAGCCAGCCGGCCCT 1620
QY 1621 GCCCAGCTTCAATTTACAGACCTGACCTTCCTCAGCTCCCTCCCTGCTGCGAGAGCTGA 1680
DB 1621 GCCCAGCTTCAATTTACAGACCTGACCTTCCTCAGCTCCCTCCCTGCTGCGAGAGCTGA 1680
QY 1681 ACATAGACTTGCACCTTGGATGTCACCTGAGTGCACATGGAGTGTATAGGAGCATCA 1740
DB 1681 ACATAGACTTGCACCTTGGATGTCACCTGAGTGCACATGGAGTGTATAGGAGCATCA 1740
QY 1741 TACCAAGGCTACTGTTGGACATGGGGGCCAAAACGCTAAACGCCACCTTCTGGAAAG 1800
DB 1741 TACCAAGGCTACTGTTGGACATGGGGGCCAAAACGCTAAACGCCACCTTCTGGAAAG 1800
QY 1801 GGAATGCAAAAGGCTTTGGGGGTGATGGAAGAAGCCTTTTACAATGATACCAATTAACCT 1860
DB 1801 GGAATGCAAAAGGCTTTGGGGGTGATGGAAGAAGCCTTTTACAATGATACCAATTAACCT 1860
QY 1861 GCCCTGAAAAAGGCATAGTGGG 1883
DB 1861 GCCCTGAAAAAGGCATAGTGGG 1883

RESULT 3
LOCUS BC027930 1951 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens death effector domain containing 2, mRNA (cDNA clone
ACCESSION BC027930
VERSION BC027930.1 GI:20379818
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1951)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F.,
Diatchenko,L., Martinina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,K., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramsen,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1951)
Strausberg,R.
Direct Submission
Submitted (08-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

```

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akheri N., Aylee K., Beckstrom-Sternberg S.M., Benjamin B., Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S., Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P., Hansen N., Ho S.L., Karling E., Kwong P., Latic P., Legaspi R., Maduro Q.L., Mastello C., Maskeri B., Mastrian S.D., McCloskey J.C., McDowell J., Pearson R., Stantripo S., Thomas P.J., Touchman J.W., Tsugeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAX Plate: 49 Row: b Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19923049.

FEATURES

source
1. 1951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:35529 IMAGE:5199772"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_1ib="NIH_MGC_114"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1951
/gene="DEDD2"
/note="Synonym: FLAME-3"
/db_xref="LOCUSID:162989"
69. 1049
/gene="DEDD2"
/codon_start=1
/product="death effector domain-containing DNA binding protein 2"
/protein_id="AAH27930.1"
/db_xref="GI:20379819"
/db_xref="LOCUSID:162989"
/translation="MALISSTPAQPMEDRECLDYGLMLSHMPFVGGSLRECELET
LAFILDEPAGAGGARARSGLLELEBRROCSNRLIGOLRLVLAARDLPHL
AKRRPVPSPERYSTSSSRTGSSCRROQSSASQOQWETSPTKRRS
RRPBGGARRRRGAPAPQOQSEAPSPSEKVCVCDRLRVRAVCEHPLEGV
SRPQALAROLDVFCATAVLRSRDLGVCDIKPSHSYLDAPFWDYLSGLLQALR
GVFLREALREAVGREAVRLVSDVADYEAQRRLMLEBEGRRPTEAS"

ORIGIN

Query Match 93.2%; Score 1754.2; DB 9; Length 1951;
Best Local Similarity 96.7%; Pred. No. 1.9e-295;
Matches 1825; Conservative 0; Mismatches 3; Indels 59; Gaps 1;
QY 56 CCGCCAGGATGATGATGAGAGCCAGAAATGCGCGGCTTGTAGCTTCCGCTCCCTC 115
DB 1 CGCCAGGATGATGAGAGCCAGAAATGCGCGGCTTGTAGCTTCCGCTCCCTC 60
QY 116 CCGCGGATGCGCTATCCGGGTGACCCCGGCCCGTGTGGAGAGAGATGAGTGC 175
DB 61 CCGCGGATGCGCTATCCGGGTGACCCCGGCCCGTGTGGAGAGAGATGAGTGC 120
QY 176 TGGACTACTAGCGGATGCTGTGCTTACCGTATGTTGAGGTGTGGCGGCAACTGA 235
DB 121 TGGACTACTAGCGGATGCTGTGCTTACCGTATGTTGAGGTGTGGCGGCAACTGA 180
QY 236 CCGAGTGCAGCTGAGAGCTTCGCTTCTGTGATAGAGCTCTGAGCGCGCGGAG 295
DB 181 CCGAGTGCAGCTGAGAGCTTCGCTTCTGTGATAGAGCTCTGAGCGCGCGGAG 240
QY 296 GCTTAGCCCGGCGCGCAGAGCCTAGAGCTCTGTGAGGTGAGAGCCCGCGGCACT 355
DB 241 GCTTAGCCCGGCGCGCAGAGCCTAGAGCTCTGTGAGGTGAGAGCCCGCGGCACT 300

QY 356 GCGCGAGAGCAACCTGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 415
 Db 301 GCGAGAGAGCAACCTGCGGCTGCTGGGGCAACCTCTGCGGCTGCTGGCCCGCCAGCACC 360
 QY 416 TGTCTGCGGCACTGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 475
 Db 361 TGTCTGCGGCACTGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 420
 QY 476 GCACTCTGCGGCACTGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 535
 Db 421 GCACTCTGCGGCACTGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 480
 QY 536 GTTCTGCAAAATTTCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 595
 Db 481 GTTCTGCAAAATTTCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 540
 QY 596 GCGGAGTGGGGGCGGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 655
 Db 541 GCGGAGTGGGGGCGGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 600
 QY 656 CACCCAGAGAGCACTGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 715
 Db 601 CACCCAGAGAGCACTGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 660
 QY 716 TCCGCTGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 775
 Db 661 TCCGCTGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 720
 QY 776 CATCCGCGGCGGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 835
 Db 721 CATCCGCGGCGGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 780
 QY 836 TGTCTGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 895
 Db 781 TGTCTGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 840
 QY 896 ATCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 955
 Db 841 ATCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 900
 QY 956 TGTCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1015
 Db 901 TGTCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 960
 QY 1016 GTGTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1075
 Db 961 GTGTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1020
 QY 1076 GGGGCGGCGGCGGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1135
 Db 1021 GGGGCGGCGGCGGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1080
 QY 1136 AGTCTCTGCGGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1176
 Db 1081 AGTCTCTGCGGCGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1140
 QY 1177 -----CCCTTGAAGAGGCGGCGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1196
 Db 1141 CTAGCATCTTTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1200
 QY 1197 CAGAGTGTGGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1256
 Db 1201 CAGAGTGTGGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1260
 QY 1257 TACCCCGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1316
 Db 1261 TACCCCGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1320
 QY 1317 GAGGCGGAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1376
 Db 1321 GAGGCGGAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1380
 QY 1377 CTGTCTGCACTGTCAAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1436

Db 1381 CTGTCTGCACTGTCAAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1440
 QY 1437 TTTCCCTGGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1496
 Db 1441 TTTCCCTGGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1500
 QY 1497 TCTCTGAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1556
 Db 1501 TCTCTGAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1560
 QY 1557 TGGAGCTGGGGTGCATGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1616
 Db 1561 TGGAGCTGGGGTGCATGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1620
 QY 1617 CCTTGGCCAGCTTCAATTTAAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1676
 Db 1621 CCTTGGCCAGCTTCAATTTAAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1680
 QY 1677 CTGACATGAGCTTGAAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1736
 Db 1681 CTGACATGAGCTTGAAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1740
 QY 1737 ATCATCAAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1796
 Db 1741 ATCATCAAGAGGCTGCTGAGAGGCTGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1800
 QY 1797 AAGGGAATGCAAGGCTTGGGGGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1856
 Db 1801 AAGGGAATGCAAGGCTTGGGGGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1860
 QY 1857 AACTGCCCTGGAAGGCAAGGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1883
 Db 1861 AACTGCCCTGGAAGGCAAGGCTGAGAGGCTGCTGGGGCACTCTGCGGCTGCTGGCCCGCCAGCACC 1887

RESULT 4
 AY125488 2012 bp mRNA linear PRI 18-SEP-2002
 LOCUS
 DEFINITION Homo sapiens death effector domain-containing DNA-binding 2
 ACCESSION AY125488
 VERSION AY125488.1 GI:22475163
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS Lee J.C., Schickling O., Stegh, A.H., Oshima, R.G., Dinsdale, D.,
 Cohen, G.M. and Peter, M.E.
 TITLE DEDD regulates degradation of intermediate filaments during
 apoptosis
 JOURNAL J. Cell Biol. 158 (6), 1051-1066 (2002)
 MEDLINE 22220288
 PUBMED 12235123
 REFERENCE
 AUTHORS Peter, M.E.
 TITLE Direct Submision
 JOURNAL Submitted (21-JUN-2002) Ben May Cancer Institute, University of
 Chicago, 924 E. 57th Street, R112, Chicago, IL 60637, USA
 FEATURES
 source
 1..2012
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..2012
 /gene="DEDD2"
 84..1049
 /gene="DEDD2"
 /codon_start=1
 /product="death effector domain-containing DNA-binding 2
 protein"

/protein_id="AA095240.1"
 /db_xref="GI:22475164"
 /translation="MALSGSTPAWPCMEDECDIDYGMLSLHMFVWGQULTELEL
 LAFLDEAGAGGLARASGLLELLERRQCDSEINRLGOLRLVLRADLLPHL
 AKRRRVPERYSYGTSKTEGSKLRRSSSSAKROSGSPPTKORSSRGP
 GGARRRRRGAAPAPPOQSEPARPSSEKXTCDIRLRVREYHEHGLAQVSRPO
 ALAROLDIVGQATVLRSDLSVSDIFPSLSLTYDARWGYEHSALLQALRVPLT
 EALREAVGEAVRLVSVDEADYEMGRRLIMEEGRRPTLAS"

Query Match 91.5%; Score 1722.2; DB 9; Length 2012;
 Best Local Similarity 95.9%; Pred. No. 6, 9e-290;
 Matches 1818; Conservative 0; Mismatches 3; Indels 74; Gaps 2;

ORIGIN

QY 48 GCGGCGCTGCGCCAGAGTATCGACGCGCAGATATGCGCGGTTCTGAGCTTGTTCG 107
 DB 8 GGGGCGCTGCGCCAGAGTATCGACGCGCAGATATGCGCGGTTCTGAGCTTGTTCG 67

QY 108 CCTCCCTCCCCGGGGAATGGCGTATCCGGGTGACCCCGGCGCTGCTGGGAGAGAGA 167
 DB 68 CCTCCCTCCCCGGGGAATGGCGTATCCGGGTGACCCCGGCGCTGCTGGGAGAGAGA 127

QY 168 TGAAGTCTGAGACTATCAAGGAGATGCTGTCGCTTACCGTATGTTGAGGTGGCGG 227
 DB 128 TGAAGTCTGAGACTATCAAGGAGATGCTGTCGCTTACCGTATGTTGAGGTGGCGG 187

QY 228 GCAACTGACCGAGTGGAGCTGCTGCGCTTCTGCTGATGAGGCTTCGCGC 287
 DB 188 GCAACTGACCGAGTGGAGCTGCTGCGCTTCTGCTGATGAGGCTTCGCGC 247

QY 288 CGCGGAGGCTTACCGCGGCGCGCGCGCTTGAAGCTTCTGCTGAGCTGAGCGCG 347
 DB 248 CGCGGAGGCTTACCGCGGCGCGCGCGCTTGAAGCTTCTGCTGAGCTGAGCGCG 307

QY 348 CGGCGAGTGGGCGAGAGCACTGCGGCTGCTGAGGCAACTCTCGCGCTGTCGCGG 407
 DB 308 CGGCGAGTGGGCGAGAGCACTGCGGCTGCTGAGGCAACTCTCGCGCTGTCGCGG 367

QY 408 CCAAGACTGCTGCGCAGCTGCGGCGCAAGCGGCGCGCGCACTGCTTCAGAACGCTA 467
 DB 368 CCAAGACTGCTGCGCAGCTGCGGCGCAAGCGGCGCGCGCACTGCTTCAGAACGCTA 427

QY 468 TGAAGTGGAGCTTCAAGCTTCAAGAGAGAGAGGAGTATGCTGCGCTGCGCGCA 527
 DB 428 TGAAGTGGAGCTTCAAGCTTCAAGAGAGAGAGGAGTATGCTGCGCTGCGCGCA 487

QY 528 GTCAAGCACTTCTGCAAAATTTCTCAGCAGGGTCACTGAGAGACAGGCTCCCGCCCAACAA 587
 DB 488 GTCAAGCACTTCTGCAAAATTTCTCAGCAGG-----GGCTCCCGCCCAACAA 532

QY 588 GCGGCAAGCGGCGAGTGGGGCGGGCCCAAGTGTGTCCAGACGCGCGGAGAGAGGCG 647
 DB 533 GCGGCAAGCGGCGAGTGGGGCGGGCCCAAGTGTGTCCAGACGCGCGGAGAGAGGCG 592

QY 648 CCAAGCGGCAAGCGGCGAGTGGGGCGGGCCCAAGCTTCTCTGAGAGGCAAAATGAC 707
 DB 593 CCAAGCGGCAAGCGGCGAGTGGGGCGGGCCCAAGCTTCTCTGAGAGGCAAAATGAC 652

QY 708 CTGTGACATCCGGCTCCGGGTTGAGAGAGAGTACTGAGAGCATGGGCGAGCTTGGAGA 767
 DB 653 CTGTGACATCCGGCTCCGGGTTGAGAGAGAGTACTGAGAGCATGGGCGAGCTTGGAGA 712

QY 768 GGGCGTGGCAATCCCGGCGGGCCCAAGGCGTGGCGCGGCACTGAGAGCTGTGGAGCG 827
 DB 713 GGGCGTGGCAATCCCGGCGGGCCCAAGGCGTGGCGCGGCACTGAGAGCTGTGGAGCG 772

QY 828 CACCGAGTGTCTGCGCTCAAGGAGCTGCGGCTTGTGATGATCAATCAAGTTCACAG 887
 DB 773 CACCGAGTGTCTGCGCTCAAGGAGCTGCGGCTTGTGATGATCAATCAAGTTCACAG 832

QY 888 GCTCTCTATCTGAGAGGCTTCTGAGGCGCACTGAGAGGCGCTGTCGAGGCGCT 947
 DB 833 GCTCTCTATCTGAGAGGCTTCTGAGGCGCACTGAGAGGCGCTGTCGAGGCGCT 892

QY 948 GCGGCGCTGCTTCTGAGTGAAGGCTTGGAGAGGCTGTGGCGCGGAGAGCTGTTCGCT 1007
 DB 893 GCGGCGCTGCTTCTGAGTGAAGGCTTGGAGAGGCTGTGGCGCGGAGAGCTGTTCGCT 952

QY 1008 GCTGGTCACTGAGTGAAGGCTGAGTATGAGGCTGGCGGCGCGCTGTTCGAGTGA 1067
 DB 953 GCTGGTCACTGAGTGAAGGCTGAGTATGAGGCTGGCGGCGCGCTGTTCGAGTGA 1012

QY 1068 GGAGGAGGCGGCGCGCGCGCGAGAGGCTTCTGATTCAGAGTGGAGGATTC 1127
 DB 1013 GGAGGAGGCGGCGCGCGCGCGAGAGGCTTCTGATTCAGAGTGGAGGATTC 1072

QY 1128 CACTTCAGAGTTCGCGGCGCACTTCTCTGGGAGAGAGCACTTCTA----- 1176
 DB 1073 CACTTCAGAGTTCGCGGCGCACTTCTCTGGGAGAGAGCACTTCTA----- 1132

QY 1177 -----CCCTTGACAGC 1188
 DB 1133 CTGTCACTGAGCACTTTGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1192

QY 1189 CCTCCACAGAGATGGGCTTGAAGGCTTAACCATTTCCAGTGAATTCCTTCCAG 1248
 DB 1193 CCTCCACAGAGATGGGCTTGAAGGCTTAACCATTTCCAGTGAATTCCTTCCAG 1252

QY 1249 ACTCTCTCAATCCCGAGGTGTGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1308
 DB 1253 ACTCTCTCAATCCCGAGGTGTGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1312

QY 1309 TCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1368
 DB 1313 TCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1372

QY 1369 ATCTTGGGCTGTCTGACATGTGACAGTGCACACATGCGTATGCTCACTGCGCTG 1428
 DB 1373 ATCTTGGGCTGTCTGACATGTGACAGTGCACACATGCGTATGCTCACTGCGCTG 1432

QY 1429 CTGAGATCTTCCCTGGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCT 1488
 DB 1433 CTGAGATCTTCCCTGGGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCTTGGCGCT 1492

QY 1489 AGGGCTTCTCTGCAAGGAGCTTGAATTTGACACACAAAGGAGGAGGAGGAGGAG 1548
 DB 1493 AGGGCTTCTCTGCAAGGAGCTTGAATTTGACACACAAAGGAGGAGGAGGAGGAG 1552

QY 1549 GTGGGCACTGAGCTGGGCTGACATGGGCGCTGCTCACTTGGCGCACTTCCAGCG 1608
 DB 1553 GTGGGCACTGAGCTGGGCTGACATGGGCGCTGCTCACTTGGCGCACTTCCAGCG 1612

QY 1609 AGCAGAGGCGCTGGCGAGCTTCAATTTAGAGCTGACTTCTCTCACTTCCCGCTGCT 1668
 DB 1613 AGCAGAGGCGCTGGCGAGCTTCAATTTAGAGCTGACTTCTCTCACTTCCCGCTGCT 1672

QY 1669 GTCCAGAGCTGAACATGACCTTGGCACTTGGAGTCACTGGAGTGCATAGGAGTGT 1728
 DB 1673 GTCCAGAGCTGAACATGACCTTGGCACTTGGAGTCACTGGAGTGCATAGGAGTGT 1732

QY 1729 ATGGAGCATATTAACCAAGGCTTATGAGTGAATGGGCGCAAAACAGTAAACAGCGAC 1788
 DB 1733 ATGGAGCATATTAACCAAGGCTTATGAGTGAATGGGCGCAAAACAGTAAACAGCGAC 1792

QY 1789 CTTCTTGAAGGAGATGCAAGGCTTGGGCTGATGAGAAAGACCTTTTACAAATGAT 1848
 DB 1793 CTTCTTGAAGGAGATGCAAGGCTTGGGCTGATGAGAAAGACCTTTTACAAATGAT 1852

QY 1849 ACCAATTAACCTGCGTGAAGGAGGATGAGG 1883
 DB 1853 ACCAATTAACCTGCGTGAAGGAGGATGAGG 1887

RESULT 5
 BC013372
 LOCUS BC013372 2005 bp mRNA linear PRI 24-FEB-2004

DEFINITION Homo sapiens death effector domain containing 2, mRNA (cDNA clone MGC:16414 IMAGE:3941907), complete cds.

ACCESSION BC013372

VERSION BC013372.2 GI:33872465

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2005)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bluetow, K.H., Scheet, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Datchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Umedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viallet, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Blakesley, R.W., Touchman, J.W., Green, E.D., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Buckson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2005)
Strausberg, R.

AUTHORS Direct Submission

TITLE Submitted (31-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15426522.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadansystemsbio.org
Madan, Stephanie Rodriguez, Amy Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Helton, Mark Kettelman, Anuradha

FEATURES
source
1..2005
location/Qualifiers
1..2005
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:16414 IMAGE:3941907"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2005
/gene="DEDD2"
/note="synonym: FLAMB-3"
/db_xref="LocusID:162989"
75..632
/gene="DEDD2"
/codon_start=1
/product="DEDD2 protein"

gene

CDS

ORIGIN

Query Match 90.4%; Score 1701.4; DB 9; Length 2005;
Best Local Similarity 95.9%; Pred. No. 2.9e-286;
Matches 1818; Conservative 0; Mismatches 1; Indels 76; Gaps 4;

/protein_id="AAH13372.2"
/db_xref="GI:33872465"
/db_xref="LocusID:162989"
/translation="MALSGSTPAFCWEEDRLDYGMLSHRMEVVGQVLTCELEL
LAFILDEPAAGLARAARSGLELLEBERGCDDBSNLRLQLRLVLAHDLPL
ARRRRPVSPEPYSYGTSSSKRTESGCRRRRSSSANSQGSPPKRRRRGRPS
GGARRRRRGAAPAPQ00SEPAQTFL"

50 GGCCGTCGCCAGAGATCATGACGACGACAAATCTGCGCGGTTCTGAGCTTTCCGCC 109
1 GGCCGTCGCCAGAGATCATGACGACGACAAATCTGCGCGGTTCTGAGCTTTCCGCC 60

110 TCCTCTCCCGGAGATGCGCTATCCGGTTCGACCCCGCCGCTGCTGGAGAGATG 169
61 TCCTCTCCCGGAGATGCGCTATCCGGTTCGACCCCGCCGCTGCTGGAGAGATG 120

170 AGTGCCTGACCTACTAGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 229
121 AGTGCCTGACCTACTAGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180

230 AACTGACGAGTGCAGAGCTGAGCTCTGAGCTTTCTGCTGATGAGAGCTCTGCGCCG 289
181 AACTGACGAGTGCAGAGCTGAGCTCTGAGCTTTCTGCTGATGAGAGCTCTGCGCCG 240

290 CCGAGGCTTACCGCCGCGCGCGCGCGCGCTGAGCTCTGCTGAGAGCTGAGCGCCG 349
241 CCGAGGCTTACCGCCGCGCGCGCGCGCGCTGAGCTCTGCTGAGAGCTGAGCGCCG 300

350 GGCAGTCCGCGAGAGAGAACTGCGGTGCTGGGCGAACTCTGCGGTGCTGCGCCGCC 409
301 GGCAGTCCGCGAGAGAGAACTGCGGTGCTGGGCGAACTCTGCGGTGCTGCGCCGCC 360

410 ACGAGCTGCTGCGCGAGCTGAGCGCGCGCGCGCGCGAGTCTCTCAGAAAGCTATA 469
361 ACGAGCTGCTGCGCGAGCTGAGCGCGCGCGCGCGCGAGTCTCTCAGAAAGCTATA 420

470 GCTATGCGACCTCCAGCTCTTCAAAAGAGACAGAGGTAAGCTGCGCGCTGCGAGT 529
421 GCTATGCGACCTCCAGCTCTTCAAAAGAGACAGAGGTAAGCTGCGCGCTGCGAGT 480

530 CAAGCAATTTCTGAAATTTCTCAGAGGCTCAGTGGAGAGAGAGTCTCCGCCCAAGC 589
481 CAAGCAATTTCTGAAATTTCTCAGAGGCTCAGTGGAGAGAGAGTCTCCGCCCAAGC 525

590 GGCAGCGCGAGAGTCCGCGCGCGCGCGCGAGTGTGTCAGACGCGCGCGAGAGGCGCC 649
526 GGCAGCGCGAGAGTCCGCGCGCGCGCGCGAGTGTGTCAGACGCGCGCGAGAGGCGCC 585

650 CAGCGGACCCGACGACGACGACGACGCGC-CAGACTTCTCTGAAGCAAGTAC 708
586 CAGCGGACCCGACGACGACGACGCGCAGACTTCTCTGAAGCAAGTAC 645

709 TGTGACATCCGCGCTCCGCGGTTTCCAGAGATATCTGCGAGATGGCCAGCTTGGACAG 768
646 TGTGACATCCGCGCTCCGCGGTTTCCAGAGATATCTGCGAGATGGCCAGCTTGGACAG 705

769 GGCCTGACATCCGCGCGCGCGCGCGCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 827
706 GGCCTGACATCCGCGCGCGCGCGCGCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 765

828 CACCGAGTGTGCTGCTCAAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 887
766 CACCGAGTGTGCTGCTCAAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 825

888 GCTCTCTATCTGAGAGCTTCTGAGGCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 947
826 GCTCTCTATCTGAGAGCTTCTGAGGCGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 885

QY 948 GCGGGGCGTGTCTGATGAGGCGCTGCGAAGGCTGTGGGCGGGAGGCTGTGGCT 1007
DB 886 GCGGGGCGTGTCTGATGAGGCGCTGCGAAGGCTGTGGGCGGGAGGCTGTGGCT 945
QY 1008 GCTGTCAGTGTGATGAGGCTGACTATGAGGCTGGCGGGCGGCTGTGTGCTGATGA 1067
DB 946 GCTGTCAGTGTGATGAGGCTGACTATGAGGCTGGCGGGCGGCTGTGTGCTGATGA 1005
QY 1068 GGAGGAAGGGGGGCGGGCGCGGACAGAGGCTTCTGATCCAGAGCTGGCAGATTGATCC 1127
DB 1006 GGAGGAAGGGGGGCGGGCGCGGACAGAGGCTTCTGATCCAGAGCTGGCAGATTGATCC 1065
QY 1128 CACTCCAGTCTCCGGGCGCACTTCTCCGGGAGAGACGACATCTCTA----- 1176
DB 1066 CACTCCAGTCTCCGGGCGCACTTCTCCGGGAGAGACGACATCTCTA----- 1125
QY 1177 -----CCCTTGAAGC 1188
DB 1126 CTGTCACTAGATCTTGAAGACTGCGACAGAGCCGGACAGAGCCCTTGAAGC 1185
QY 1189 CCCTCCACAGATGTGGGCTGTGAGGCTTAAACATTTTCAGCTGATTTCTTCCAG 1248
DB 1186 CCCTCCACAGATGTGGGCTGTGAGGCTTAAACATTTTCAGCTGATTTCTTCCAG 1245
QY 1249 ACTCTCTTCAATCCCGCAGTGTGCGCTTACCTCCGGAGGCGGGGCTGGGCTGTATC 1308
DB 1246 ACTCTCTTCAATCCCGCAGTGTGCGCTTACCTCCGGAGGCGGGGCTGGGCTGTATC 1305
QY 1309 TCAGAAAGGAGGGGCGAGCTACACACTACAAAGGCGCCCTGACATTTGATCTCTG 1368
DB 1306 TCAGAAAGGAGGGGCGAGCTACACACTACAAAGGCGCCCTGACATTTGATCTCTG 1365
QY 1369 ATCTTGGGCTGTGTGCACTGTCAAGGTGACACACTGCTCATGTCTACACTGCCCC 1428
DB 1366 ATCTTGGGCTGTGTGCACTGTCAAGGTGACACACTGCTCATGTCTACACTGCCCC 1425
QY 1429 CTGAGATCTTCCCGGCGCTTCCCTGCGCTGCTCCGAGCACAACCTTTTGGCTTA 1488
DB 1426 CTGAGATCTTCCCGGCGCTTCCCTGCGCTGCTCCGAGCACAACCTTTTGGCTTA 1485
QY 1489 AGGCTTCTCTCTCAGGACTCTTAATTGACCAACAACCTGGGCTTACGACATCA 1548
DB 1486 AGGCTTCTCTCTCAGGACTCTTAATTGACCAACAACCTGGGCTTACGACATCA 1545
QY 1549 GTGGGCACTGGAGCTGGGCTGCAATGGGCTGTCTCACTTGGCCACACATCTCAAGCC 1608
DB 1546 GTGGGCACTGGAGCTGGGCTGCAATGGGCTGTCTCACTTGGCCACACATCTCAAGCC 1605
QY 1609 AGGCAAGGCGGCGGCGGCTTCAATTTACAGACTGATCTCTCACTTCCCGGCTGCT 1668
DB 1606 AGGCAAGGCGGCGGCGGCTTCAATTTACAGACTGATCTCTCACTTCCCGGCTGCT 1665
QY 1669 GTCCAGAGCTGACATGACCTTCACTTGGATGTCACTTGAAGTGTCAATGGAGTGT 1728
DB 1666 GTCCAGAGCTGACATGACCTTCACTTGGATGTCACTTGAAGTGTCAATGGAGTGT 1725
QY 1729 ATGGCAGCATCATACCAAGGCTTACTGTGACATGGGCGCAAAACAGTAAACAGCCAC 1788
DB 1726 ATGGCAGCATCATACCAAGGCTTACTGTGACATGGGCGCAAAACAGTAAACAGCCAC 1785
QY 1789 CTCTTGGAAAGGAAGGAAGGCTTGGGGGTGATGGAAGAAACCTTTTAAATGAT 1848
DB 1786 CTCTTGGAAAGGAAGGAAGGCTTGGGGGTGATGGAAGAAACCTTTTAAATGAT 1845
QY 1849 ACCAATTAAGTCCCTGGAAGAGGCTAGGTGG 1883
DB 1846 ACCAATTAAGTCCCTGGAAGAGGCTAGGTGG 1880

RESULT 6
AX431308 1924 bp DNA linear PAT 28-JUN-2002
LOCUS AX431308
DEFINITION Sequence 17 from Patent WO0240680.

ACCESSION AX431308
VERSION AX431308.1 GI:21656177
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S.H., Reed, J.C.,
Roth, W. and Stenner-Liwen, F.
TITLE Novel death domain proteins
JOURNAL Patent: WO 0240680-A 17 23-MAY-2002;
BURNHAM INST (US)
FEATURES
source Location/Qualifiers
1..1924
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
91..>1044
/note="unamed protein product"
/codon_start=1
/protein_id="CAD37280.1"
/db_xref="GI:21656178"
/translation="MALSGTPAPCMEDDECLDYGYMLSLHMFVYGGQLTRECETL
IAFLIDARAGAGIARASGLLELLEBRCODESNRLIGOLLRVLRBDLPHL
AKRRRPVSPERYSTSSSKRTGSCRRROSSANSOQOMETSPPTKRRRS
RRPFGAARRRRGAAPAVQOQSEPARSBESKCDILRLRAYCEHPELQGV
SRPOLAROLDVFOATVALRSRDLGVCIKFSELSYDAFMGDYLSGALLQALR
GVFLTEALREAVREAVRLVSVDBADYBAGRRLILMEEG"

ORIGIN
Query Match 89.3%; Score 1681.4; DB 6; Length 1924;
Best Local Similarity 95.5%; Pred. No. 8.9e-283;
Matches 1828; Conservative 0; Mismatches 19; Indels 67; Gaps 7;
QY 33 CCNAAACAGTGGCTGCGGCGCGGCGGAGTCACTGCGAGCAATCTGCGCGGT 92
DB 1 CCNAAACAGTGGCTGCGGCGCGGCGGAGTCACTGCGAGCAATCTGCGCGGT 59
QY 93 TCTGAGCTTGTTCGCGCTCTCCCGCGGAAATGGCGCTATCGGGTCGACCCCGGCCC 152
DB 60 TCTGAGCTTGTTCGCGCTCTCCCGCGGAAATGGCGCTATCGGGTCGACCCCGGCCC 119
QY 153 GTGCTGGAGAGAGATGATGCTGCTGATCTAATCGGAGATGCTGTGCTTCACTGATTT 212
DB 120 GTGCTGGAGAGAGATGATGCTGCTGATCTAATCGGAGATGCTGTGCTTCACTGATTT 179
QY 213 CGAGGTGTGGGGGCGGCACTGACCGAGTGCAGAGTGCAGTCTCTGCGCTTCTGCTGGA 272
DB 180 CGAGGTGTGGGGGCGGCACTGACCGAGTGCAGAGTGCAGTCTCTGCGCTTCTGCTGGA 239
QY 273 TGAAGCTCTGTGGCGCGCGGAGGCTTGAAGCCCGGCGCGGCGGCTTGAAGCTCTGTGCT 332
DB 240 TGAAGCTCTGTGGCGCGCGGAGGCTTGAAGCCCGGCGCGGCGGCTTGAAGCTCTGTGCT 299
QY 333 GGAAGCTGAGCGCGCGGCGGAGTGCAGAGCAACCTGCGGCTGTCTGGGCAACTCTCT 392
DB 300 GGAAGCTGAGCGCGCGGCGGAGTGCAGAGCAACCTGCGGCTGTCTGGGCAACTCTCT 359
QY 393 GGGCGTGTGGCGCGCGGCGGAGCTGTGCGGACCTGCGGCGGCGGCGGCGGCGGAGT 452
DB 360 GGGCGTGTGGCGCGCGGCGGAGCTGTGCGGACCTGCGGCGGCGGCGGCGGCGGAGT 419
QY 453 GTCTCCAGAACGTATAGCTATGCACTTCAAGCTTTCAAAAGAGACAGAGGTAGCTG 512
DB 420 GTCTCCAGAACGTATAGCTATGCACTTCAAGCTTTCAAAAGAGACAGAGGTAGCTG 479
QY 513 CGGTGCGCGTGGCGAGTCAAGCACTTTCGAAATTTTCAGCAGGCTCACTGGAGACAG 572
DB 480 CGGTGCGCGTGGCGAGTCAAGCACTTTCGAAATTTTCAGCAGGCTCACTGGAGACAG 539
QY 573 CTCGCCCGCAACAGCGGCGAGCGGCGGAGTGGGGCGGCGGCGGAGTGTGTCCAGAG 632

```

Db      540 CTCCTCCCCCAACGAGCGGAGCGGAGTGGGGCCGCGCCAGTGGTGTCCAGACG 599
Qy      633 GCGGCGAGAGGGGCCCCAGCGGAGCCCGCAGCAGTGTAGAGGCCCGCCAGACCTTCTC 692
Db      600 GCGGCGAGAGGGGCCCCAGCGGAGCCCGCAGCAGTGTAGAGGCCCGCCAGACCTTCTC 659
Qy      693 TGAAGGCAAAAGTGAAGTGTGATCCGCGCTCCGGGTTTGAGACAGATCTGCGAGCATGG 752
Db      660 TGAAGGCAAAAGTGAAGTGTGATCCGCGCTCCGGGTTTGAGACAGATCTGCGAGCATGG 719
Qy      753 GCCAGCCTTGGAGAGGGGCGTGGGATCCCGCGGCGCCAGGCGGTGGCGCGGAGCTGGA 812
Db      720 GCCAGCCTTGGAGAGGGGCGTGGGATCCCGCGGCGCCAGGCGGTGGCGCGGAGCTGGA 779
Qy      813 CGTGTGTGGGAGGAGCCAGCGGAGTGTGCGCTCAAGGGAAGCTTGGGCTCTGTGTGTGGA 872
Db      780 CGTGTGTGGGAGGAGCCAGCGGAGTGTGCGCTCAAGGGAAGCTTGGGCTCTGTGTGTGGA 839
Qy      873 CATCAAGTTCTCAGAGCTCTCTATCTGTGAGCGCCTTCTGGGGGAGTAACTGAGTGGCGC 932
Db      840 CATCAAGTTCTCAGAGCTCTCTATCTGTGAGCGCCTTCTGGGGGAGTAACTGAGTGGCGC 899
Qy      933 CTGTGCTGAGAGCCCTTGGGGGCGGTGTCTGTGAGAGGCCCTGGAGAGGCGTGTGGGCGG 992
Db      900 CTGTGCTGAGAGCCCTTGGGGGCGGTGTCTGTGAGAGGCCCTGGAGAGGCGTGTGGGCGG 959
Qy      993 GGAAGGCTGTTCGCTGTGTGATGTGATGAGGCTGATATAGAGGCTGGCGGCGCG 1052
Db      960 GGAAGGCTGTTCGCTGTGTGATGTGATGAGGCTGATATAGAGGCTGGCGGCGCG 1019
Qy      1053 CTTGTGCTGATGAGAGGAGGAGGGGGCGGCGCCGACAGAGGCGCTCTGTGATCAGAGAC 1112
Db      1020 CTTGTGCTGATGAGAGGAGGAGGGGGCGGCGCCGACAGAGGCGCTCTGTGATCAGAGAC 1079
Qy      1113 TGGGAGATTTGATATCCCACTTCCAGAGTCTCCGGGAGCACTTCTCTGGGAGGAGCAACCATC 1172
Db      1080 TGGGAGATTTGATATCCCACTTCCAGAGTCTCCGGGAGCACTTCTCTGGGAGGAGCAACCATC 1139
Qy      1173 TCTA----- 1176
Db      1140 TCTA----- 1199
Qy      1177 ---CCCTTGAAGAGCCCTTCCAGAGAGTGTGGCTGTGAGGCTTAAACATTTTCAAGCT 1233
Db      1200 AGGCGCCCTTGAAGAGCCCTTCCAGAGAGTGTGGCTGTGAGGCTTAAACATTTTCAAGCT 1259
Qy      1234 GAGTTTCTTCTCCAGACTCTCTCAACCCCGAGGTGTGCGCTTAAAGCTCGAGGGGGG 1293
Db      1260 GAGTTTCTTCTCCAGACTCTCTCAACCCCGAGGTGTGCGCTTAAAGCTCGAGGGGGG 1318
Qy      1294 GGGTGGGCTGTATCTCAGAAAGGAGGGGAGCAGCTACACCTACCAAAAGGCGCCCTTG 1353
Db      1319 GGGTGGGCTGTATCTCAGAAAGGAGGGGAGCAGCTACACCTACCAAAAGGCGCCCTTG 1378
Qy      1354 CACATTTATCTGTATCTTGGGCTGTCTGTGACTGTGACAGTGTGACACACTCGCTCATG 1413
Db      1379 CACATTTATCTGTATCTTGGGCTGTGTGTGACTGTGACAGTGTGACACACTCGCTCATG 1438
Qy      1414 CTCACTACGCGCCCTGCTGAGATCTTCCCTGGGCTCTGCGCTGGCTGTCTTCCAGACAA 1473
Db      1439 CTCACTACGCGCCCTGCTGAGATCTT-CTTGGGCTCTGCGCTGGCTGTG-TTCCAGACAA 1496
Qy      1474 CACTTTCTTGGGCTTAAAGGGCTTCTCTCAGGAGCTCTAATTATTTAGCACAA---CGAAC 1530
Db      1497 CACTTTCTTGGGCTTAAAGGGCTTCTTCTCAGGAGCTTAAATTATTTACCAACCAACCAACT 1556
Qy      1531 TGGGCTTCAAGC-AACATCAGTGGGCACTGTGAGCTGTGGGTGACATGTGGGCTGTCTCACT 1589
Db      1557 GGGGTTTCAAGCAAAATCAGTGGGCACTGTGAGCTGTGGGTGACATGTGGGCTGTCTCACT 1616
Qy      1590 TGGCCACACATCTCAGCAGCGGCGGCGCTGGCCAGCTTCAATTTACAGACTGAGCTCT 1649
Db      1617 TGGCCACAAATTTCCAGCAGCGGCGGCGCTGGCCAGCTTCAATTTACAGACTGAGCTMT 1676

```

```

Qy      1650 CTTACCTTCCCTCTGCTGTCCAGAGCTGAACATAGACTTGTGACTGTGATGTCACCTGG 1709
Db      1677 CTTACCTTCCCTCTGCTGTCCAGAGCTGAACATAGACTTGTGACTGTGATGTCACCTGG 1736
Qy      1710 AGTGTCAATGAGAGTGTATGTGCAAGATCATATACCAAGGCTTACTGTGTCATGAGGGCC 1769
Db      1737 AGTGTCAATGAGAGTGTATGTGCAAGATCATATACCAAGGCTTACTGTGTCATGAGGGCC 1796
Qy      1770 AAAACAGTAAACAGCAGCCTTCTTGGAAAGGAAATGCAAAAGGCTTGGGGGTGATGGA 1829
Db      1797 AAAACAGTAAACAGCAGCCTTCTTGGAAAGGAAATGCAAAAGGCTTGGGGGTGATGGA 1856
Qy      1830 AAGACCTTTTAAATATATCAATTAATCAATGAGGCTTGGAAAGGAGCATAGTGGG 1883
Db      1857 AAGACCTTTTAAATATATCAATTAATCAATGAGGCTTGGAAAGGAGCATAGTGGG 1910

RESULT 7
AKI30203
LOCUS      Homo sapiens cDNA FLJ26693 fis, clone MPO7953.
DEFINITION
ACCESSION AKI30203
VERSION   AKI30203.1 GI:34526957
KEYWORDS  oligo capping; fis (full insert sequence).
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1
AUTHORS   Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
          Fukuzumi,Y., Fujimori,Y., Komiyama,M., Suzuki,Y., Hata,H.,
          Nkagawa,K., Mizuno,S., Morinaga,M., Kawamura,M., Sugiyama,T.,
          Irie,R., Otsubi,T., Sato,H., Nishikawa,T., Sugiyama,A.,
          Kawakami,B., Nagai,K., Isegai,T. and Sugano,S.
          NEDO human cDNA sequencing project
          Unpublished
          2 (bases 1 to 1659)
REFERENCE 2
AUTHORS   Sugano,S. and Suzuki,Y.
TITLE     Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,
          University of Tokyo, Laboratory of Genome Structure, Human Genome
          Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
          (E-mail:elcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
          Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
          Economy, Trade and Industry of Japan; cDNA full insert sequencing;
          Research Association for Biotechnology (RAB); cDNA library
          construction and 5'-end one pass sequencing; Institute of Medical
          Science, University of Tokyo; Laboratory of Genome Structure, Human
          Genome Center; 3'-end one pass sequencing; RAB; clone selection for
          full insert sequencing; RAB and Helix Research Institute.
FEATURES
source
1. 1659
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="MPO7953"
   /cell_type="macrophage"
   /clone_1ib="MPG"
   /note="Cloning vector: pMB18SFLJ3-differentiated by
         oxidant"

ORIGIN
Query Match      66.7%; Score 1256.8; DB 9; Length 1659;
Best Local Similarity 89.0%; Pred. No. 7,5e-209;
Matches 1447; Conservative 0; Mismatches 22; Indels 157; Gaps 2;

Qy      415 CTGCTGCGGACCTGCGGCGGAGCGGCGGCGGAGTGTCTCAGAAAGCTTATAGCTAT 474
Db      33  GGGCGGTGGCGCCAGAGACTATCGGACCCCAAAATCTGTGTCTCCAGAAAGCTTATAGCTAT 92
Qy      475 GGCACCTCCAGCTTCTTCAAAAGAGACAGAGGATAGCTGCCGTGCGCAGTCAAGC 534

```



```

Db      93  GGAACCTCCAGCTCTTCAAGAGAGACAGAGGTAGTGCCTGCGCGCTGCGAGTAAAG 152
Qy      535  AGTTTGCAAAATTTCTCAGCAGAGGTCACTGGAGAGACAGAGCTCCCCCAACCAAGCGGAG 594
Db      153  AGTTTGCAAAATTTCTCAGCAGAGGTCACTGGAGAGACAGAGCTCCCCCAACCAAGCGGAG 212
Qy      595  CGGCGAGAGTGGGGCCGGCCAGTGTGTGCAGACGCGCGGAGAGAGGGGCCAGGCC 654
Db      213  CGGCGAGAGTGGGGCCGGCCAGTGTGTGCAGACGCGCGGAGAGAGGGGCCAGGCC 272
Qy      655  GCACCCCAGAGAGTCAAGAGCCCGCAGACTTCTCTGAAAGCAAAAGTGAAGTGAAGT 710
Db      273  GCACCCCAGAGAGTCAAGAGCCCGCAGACTTCTCTGAAAGCAAAAGTGAAGTGAAGT 332
Qy      711  ----- 710
Db      333  TGTACAAGAGAGGTGCCAGATCTGCTTCCGTTGAGAGACCTCCGGAAGCTTCCATTTC 392
Qy      711  -----TGAATCCGGCTTCCGGTTTGAAGCAG 736
Db      393  ATGTGGAAGGCCAAGGGAGAGAGGCTTGTCAATGACATCCGGCTCCGGGTTTGAAGCAG 452
Qy      737  AGTACTGCGAGCATGGGGCCAGCTTGTGAGAGAGGGCGTGGATCCCGCGGCCAGAGCGC 796
Db      453  AGTACTGCGAGCATGGGGCCAGCTTGTGAGAGAGGGCGTGGATCCCGCGGCCAGAGCGC 512
Qy      797  TGGCGCGGAGCTGAGAGTGTGTTGGGAGAGCACCAGAGTCTGCGCTCAAGGAGACTGG 856
Db      513  TGGCGCGGAGCTGAGAGTGTGTTGGGAGAGCACCAGAGTCTGCGCTCAAGGAGACTGG 572
Qy      857  GCTCTGTGTTTGTGACATCAAGTCTCAGAGCTCTTCTATCTGAAGCGCTTCTGGGGCG 916
Db      573  GCTCTGTGTTTGTGACATCAAGTCTCAGAGCTCTTCTATCTGAAGCGCTTCTGGGGCG 632
Qy      917  ACTACCTGATGGGCGCCCTGCTCAGAGCCCTGCGGGCGGTTCTTGACCTGAAGCCCTGC 976
Db      633  ACTACCTGATGGGCGCCCTGCTCAGAGCCCTGCGGGCGGTTCTTGACCTGAAGCCCTGC 692
Qy      977  GAGAGGCTGTGGGCGGGAGGCTGTTGCTGCTGTGATGATGATGAGGCTGACTATG 1036
Db      693  GAGAGGCTGTGGGCGGGAGGCTGTTGCTGCTGTGATGATGATGAGGCTGACTATG 752
Qy      1037  AGGCTGGCGCGGCGCCGCTGTGCTGATGAGAGAGAGAGAGGGGCGCGCCGACAGAGG 1096
Db      753  AGGCTGGCGCGGCGCCGCTGTGCTGATGAGAGAGAGAGAGGGGCGCGCCGACAGAGG 812
Qy      1097  CCTCTGATCCAGAGCTGGAGAGATTGATCCCACTTCCAGTCTCCGGGCGACCTTCTCC 1156
Db      813  CCTCTGATCCAGAGCTGGAGAGATTGATCCCACTTCCAGTCTCCGGGCGACCTTCTCC 872
Qy      1157  TGGGAGAGCAGCATTCTTA----- 1176
Db      873  TGGGAGAGCAGCATTCTTA----- 1176
Qy      1177  -----CCCTTGAAGAGCCCTTCCAGAGAGTGTGGCTCTGAAGCC 1217
Db      933  ACAGAGACCGGAGACAGAGCCCTTGAAGAGCCCTTCCAGAGAGTGTGGCTCTGAAGCC 992
Qy      1218  TAAACATTTTCCAGCTGATTTCTTCCAGAGCTCTTCCAGAGAGTGTGGCTCTGAAGCC 1277
Db      993  TAAACATTTTCCAGCTGATTTCTTCCAGAGCTCTTCCAGAGAGTGTGGCTCTGAAGCC 1052
Qy      1278  AGGCTCGGAGGCGGGGCTGTGCTGTATCTCAGAGAGGAGGAGCAGACTACACTC 1337
Db      1053  AGGCTCGGAGGCGGGGCTGTGCTGTATCTCAGAGAGGAGGAGCAGACTACACTC 1112
Qy      1338  ACCAAAGGCGCCCTGACATTTGATCTGTATCTTGGGCTGTCTGACATCTCAAGTG 1397
Db      1113  ACCAAAGGCGCCCTGACATTTGATCTGTATCTTGGGCTGTCTGACATCTCAAGTG 1172
Qy      1398  CACACATCGCTATGCTACACATGCCCCCTGCTGAAGTCTTCCCTGGGCTTGCCTTG 1457

```

```

Db      1173  CACACATGCTGATGCTACACTGCCCCCTGCTGAGATCTTCCCTGGGCTCTGCCCTGG 1232
Qy      1458  CCTGCTTCCAGACACACACTTCTTTGGCCTTAAGGGCTTCTCTCAGACCTTAATTGG 1517
Db      1233  CCTGCTTCCAGACACACTTCTTTGGCCTTAAGGGCTTCTCTCAGACCTTAATTGG 1292
Qy      1518  ACCAACAACCACTTGGGCTTTCAGCCCATCACTAGTGGGCTATGAGAGCTGGGGTGCATGG 1577
Db      1293  ACCAACAACCACTTGGGCTTTCAGCCCATCACTAGTGGGCTATGAGAGCTGGGGTGCATGG 1352
Qy      1578  GCGTGTCACTTTGCCCCACACATCTCAGCCAGCCAGGAGCCCTGCCCCAGCTTAATTAC 1637
Db      1353  GCGTGTCACTTTGCCCCACACATCTCAGCCAGCCAGGAGCCCTGCCCCAGCTTAATTAC 1412
Qy      1638  AGACTGACTCTTCTCACTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1697
Db      1413  AGACTGACTCTTCTCACTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1472
Qy      1698  GATGTCACTTGGAGTGCATGAGAGTGTATGAGAGATCATACCAAGGCTTACTGTT 1757
Db      1473  GATGTCACTTGGAGTGCATGAGAGTGTATGAGAGATCATACCAAGGCTTACTGTT 1532
Qy      1758  GCACATGGGGCCAAAACAGTAAACAGCCACTTCTTGAAGAGGAATGCAAGGCTTTG 1817
Db      1533  GCACATGGGGCCAAAACAGTAAACAGCCACTTCTTGAAGAGGAATGCAAGGCTTTG 1592
Qy      1818  GGGGTGATGAGAAAAGACTTTTACAATATGATACCAATTAACTGCCCTGGAAGGGCATTA 1877
Db      1593  GGGGTGATGAGAAAAGACTTTTACAATATGATACCAATTAACTGCCCTGGAAGGGCATTA 1652
Qy      1878  GGTGGG 1883
Db      1653  GGTGGG 1658

RESULT 8
AX322754      1230 bp      DNA      linear      PAT 07-JAN-2002
LOCUS
DEFINITION
AX322754      Sequence 4 from Patent WO0192527.
ACCESSION
AX322754.1      GI:18093742
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1      Tang Y.T., Azimzai, Y., Yue H., Burford, N., Ding, L., Elliott, V.S.,
Patterson, C., and Baugim, M.R.
Regulators of apoptosis
Patent: WO 0192527-A 4 06-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1230
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3102521CB1"

ORIGIN
Query Match      62.7%; Score 1180.8; DB 6; Length 1230;
Best Local Similarity 99.8%; Pred. No. 1,3e-195;
Matches 1182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1      AGGCGTAATTAATGAGAGGTGCCAAGATGATCAAAACAGAGTGGCTGCGGCGGCGCC 60
Db      4      AGGCGTAATTAATGAGAGGTGCCAAGATGATCAAAACAGAGTGGCTGCGGCGGCGCC 63
Qy      61      AGAGTCAATCGAGCGCAGATCTGGCGGGTTCTGAGCTTGTTCGCGCTCCCTCCCG 120
Db      64      AGGAGTCAATCGAGCGCAGATCTGGCGGGTTCTGAGCTTGTTCGCGCTCCCTCCCG 123
Qy      121      GGAATGGCGCTATCCGGGTGACCCCGGCCCTGCTGGAGAGATGATGCTTGAC 180

```


Db 124 GGAATGCGCTATCCGGGTGACACCCCGCCCGCTGAGGAGAGATGAGTGCCTGAC 183
Qy 181 TACTACGGGATGCTGTGCTTCAACCGATGTTGCAAGGTGTGGCGGGCAATGACCGAG 240
Db 184 TACTACGGGATGCTGTGCTTCAACCGATGTTGCAAGGTGTGGCGGGCAATGACCGAG 243
Qy 241 TGGAGCTGAGAGCTCTGAGCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTA 300
Db 244 TGGAGCTGAGAGCTCTGAGCTTCTGCTGATGAGGCTCTGCGCGCGCGAGGCTTA 303
Qy 301 GCCCGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGGCTGAGCGCGCGGCACTGCGC 360
Db 304 GCCCGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGGCTGAGCGCGCGGCACTGCGC 363
Qy 361 GAGAGCAACTGCGGCTGTGGGCACTCTGCGCGCTGCTGCGCGCGCGAGACTGCTG 420
Db 364 GAGAGCAACTGCGGCTGTGGGCACTCTGCGCGCTGCTGCGCGCGCGAGACTGCTG 423
Qy 421 CCGCACTGCGCGCAAGCGGCGCGGCACTGCTGCTGAGAGCTTATGAGCACC 480
Db 424 CCGCACTGCGCGCAAGCGGCGCGGCACTGCTGCTGAGAGCTTATGAGCACC 483
Qy 481 TCCAGCTCTTCAAGAGAGAGAGAGGAGCTGCGCGCTGCGCACTGAGCAGTTCT 540
Db 484 TCCAGCTCTTCAAGAGAGAGAGAGGAGCTGCGCGCTGCGCACTGAGCAGTTCT 543
Qy 541 GCAAAATCTCAGCAGGCTCAGTGGGAGACAGGCTCCCCCAACCAAGCGGCAAGCGG 600
Db 544 GCAAAATCTCAGCAGGCTCAGTGGGAGACAGGCTCCCCCAACCAAGCGGCAAGCGG 603
Qy 601 AGTCGGGGCCCGCCAGTGTGTGTCAGAGCGCGCGGAGAGGGGCCCGAGCGCAACC 660
Db 604 AGTCGGGGCCCGCCAGTGTGTGTCAGAGCGCGCGGAGAGGGGCCCGAGCGCAACC 663
Qy 661 CAGCAGAGCTCAGAGCCCGCAGACCTTCTCTGAGGCAAAAGTGAAGCTGATCCGG 720
Db 664 CAGCAGAGCTCAGAGCCCGCAGACCTTCTCTGAGGCAAAAGTGAAGCTGATCCGG 723
Qy 721 CTCGGGTTGAGAGAGTACTGAGACATGAGGCACTTGGAGAGAGGAGTGGCATCC 780
Db 724 CTCGGGTTGAGAGAGTACTGAGACATGAGGCACTTGGAGAGAGGAGTGGCATCC 783
Qy 781 CCGCGGCGCCAGGCGCTGCGCGGCACTGAGCGTGTGGGCGAGCGCAGTGTG 840
Db 784 CCGCGGCGCCAGGCGCTGCGCGGCACTGAGCGTGTGGGCGAGCGCAGTGTG 843
Qy 841 CGCTCAAGGAGACTGAGGCTGTGTGTTGATCAATCAAGTTCTGAGAGCTTCTATCTG 900
Db 844 CGCTCAAGGAGACTGAGGCTGTGTGTTGATCAATCAAGTTCTGAGAGCTTCTATCTG 903
Qy 901 GAGCGCTTCTGAGGCGACTACTGAGTGGCGGCTGCTGCAAGGCGCTGCGGCGGTTC 960
Db 904 GAGCGCTTCTGAGGCGACTACTGAGTGGCGGCTGCTGCAAGGCGCTGCGGCGGTTC 963
Qy 961 CTGAGCTAGAGCCCTGCGAGAGGCTGTGGGCGGAGAGGCTGTGCTGCTGCTGCTG 1020
Db 964 CTGAGCTAGAGCCCTGCGAGAGGCTGTGGGCGGAGAGGCTGTGCTGCTGCTGCTG 1023
Qy 1021 GATGAGCTGACTATGAGGCTGCGCGGCGGCTGTTGCTGATGAGAGAGAGAGGAGG 1080
Db 1024 GATGAGCTGACTATGAGGCTGCGCGGCGGCTGTTGCTGATGAGAGAGAGAGGAGG 1083
Qy 1081 CCGGCGCGGAGAGGCGCTCTGATCCAGAGCTGGCAGAGATGATCCCACTCCAGTCT 1140
Db 1084 CCGGCGCGGAGAGGCGCTCTGATCCAGAGCTGGCAGAGATGATCCCACTCCAGTCT 1143
Qy 1141 CCGGCGCGGAGAGGCTCTGATCCAGAGCTGGCAGAGATGATCCCACTCCAGTCT 1184
Db 1144 CCGGCGCGGAGAGGCTCTGATCCAGAGCTGGCAGAGATGATCCCACTCCAGTCT 1187

AC010247/c
LOCUS AC010247 95663 bp DNA linear PRI 13-JUL-2002
DEFINITION Homo sapiens chromosome 19 clone CTC-378H22, complete sequence.
AC010247
ACCESSION AC010247
VERSION AC010247.9 GI:21743752
KEYWORDS HTG.
SOURCE
ORGANISM Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
2 (bases 1 to 95663)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 95663)
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 95663)
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 13, 2002 this sequence version replaced gi:21637454.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
NOTE: Small insert shatter library only 85964-86181. 226bp single subclone 9479-9697.
Location/Qualifiers
1. 95663
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-378H22"
join(9479..9697,85964..86181)
/note="NOTE: Small insert shatter library only 85964-86181. 226bp single subclone 9479-9697."
ORIGIN
Query Match 58.6%; Score 1103; DB 9; Length 95663;
Best Local Similarity 95.2%; Pred. No. 2e-182;
Matches 1172; Conservative 0; Mismatches 0; Indels 59; Gaps 1;
Qy 712 GACATCCGCGCTCCGGGTTGAGAGAGTACTGAGAGTGGAGGAGGCTTGGAGAGGAGG 771
Db 90884 GACATCCGCGCTCCGGGTTGAGAGAGTACTGAGAGTGGAGGAGGCTTGGAGAGGAGG 90925
Qy 772 GTGGCATCCCGCGGCGCCCAAGCGGCTGCGCGGCACTGAGCGTGTGGGCGAGCCACC 831
Db 90924 GTGGCATCCCGCGGCGCCCAAGCGGCTGCGCGGCACTGAGCGTGTGGGCGAGCCACC 90865
Qy 832 GCACTGTGCGCTCAAGGAGCTGGGCTGTGTGTTGATCAATCAAGTTCTGAGAGCTC 891
Db 90864 GCACTGTGCGCTCAAGGAGCTGGGCTGTGTGTTGATCAATCAAGTTCTGAGAGCTC 90805
Qy 892 TCTATCTGAGAGCTTCTGAGGCGACTACTGAGTGGCGGCTGCTGAGGCGCTTGGCGG 951
Db 90804 TCTATCTGAGAGCTTCTGAGGCGACTACTGAGTGGCGGCTGCTGAGGCGGCTTGGCGG 90745
Qy 952 GCGGTGTTCTGACTGAGGCGCTTGGCGAGAGGCTGTGGGCGGAGAGGCTGTGCTGCTG 1011
Db 90744 GCGGTGTTCTGACTGAGGCGCTTGGCGAGAGGCTGTGGGCGGAGAGGCTGTGCTGCTG 90685

QY	1012	GTCAGTGTGATGAGGCTGTGACTGTAGAGCTGCGGCCGCCGCCCTGTGTGCTGTATGAGAGAG	1071
Db	90684	GTCAGTGTGATGAGGCTGTGACTGTAGAGCTGCGGCCGCCGCCCTGTGTGCTGTATGAGAGAG	90625
QY	1072	GAAAGGGGGGGGGGGCCCGACAGAGAGGCGCTCTGTATTCAGAGACTGGACAGATGTATCCACCC	1131
Db	90624	GAAAGGGGGGGGGGGCCCGACAGAGGCGCTCTGTATTCAGAGACTGGACAGATGTATCCACCC	90585
QY	1132	TCCAAGTCTCCGGGCGACCTTCTCTGGAGAGAGCACTTCTTA-----	1176
Db	90564	TCCAAGTCTCCGGGCGACCTTCTCTGGAGAGAGCACTTCTTA-----	90505
QY	1177	-----CCCCCTTGACAGCCCT	1192
Db	90504	CACCTTAGCATCTTTGAGAGACTGCGACAGGACCGGACAGAGAGGCCCCCTTGACAGCCCT	90445
QY	1193	CCCAACAGAGATGTGGGCTCTGAGGCGCTTAACCATTTTCAGCTGAGTTTCTTCCACAGACTC	1252
Db	90444	CCCAACAGAGATGTGGGCTCTGAGGCGCTTAACCATTTTCAGCTGAGTTTCTTCCACAGACTC	90385
QY	1253	CTCTTACCCCGAGGTGTGCCCTTTAGCCTCCGAGGCGGGGGCTGGGCTGTATCTGAG	1312
Db	90384	CTCTTACCCCGAGGTGTGCCCTTTAGCCTCCGAGGCGGGGGCTGGGCTGTATCTGAG	90325
QY	1313	AAGGAGAGGGGCGACAGCTACACATCCACCAAGGCGCCCGCTGCACTTTGATCTGTGATCT	1372
Db	90324	AAGGAGAGGGGCGACAGCTACACATCCACCAAGGCGCCCGCTGCACTTTGATCTGTGATCT	90285
QY	1373	TGGGCTGTCTGCACTGTGCACAGGTGCACACACTGTGCATGTCTACACTGCCCTGTCTGA	1432
Db	90264	TGGGCTGTCTGCACTGTGCACAGGTGCACACACTGTGCATGTCTACACTGCCCTGTCTGA	90205
QY	1433	GATCTTCCCTGGGCTCTGCCCTGTGCTGTTCCACAGACACACTTCTTTGGCTTAAGG	1492
Db	90204	GATCTTCCCTGGGCTCTGCCCTGTGCTGTTCCACAGACACACTTCTTTGGCTTAAGG	90145
QY	1493	CTTCTCTCTAGAGACTCTTAATTTGACCAACCAACCACTGGGCTTGAGGACATCAGTGG	1552
Db	90144	CTTCTCTCTAGAGACTCTTAATTTGACCAACCAACCACTGGGCTTGAGGACATCAGTGG	90085
QY	1553	GCACTGAGAGCTGGGGTGCACATGGGGGCTGCTACCTTGGCCACACATCTCCAGCCAGCC	1612
Db	90084	GCACTGAGAGCTGGGGTGCACATGGGGGCTGCTACCTTGGCCACACATCTCCAGCCAGCC	90025
QY	1613	AGGGCCCTGCGCCAGCTTCATTTACAGACTGTCTCTCACTTCCCCCTGCTGTCC	1672
Db	90024	AGGGCCCTGCGCCAGCTTCATTTACAGACTGTCTCTCACTTCCCCCTGCTGTCC	89965
QY	1673	AGAGCTGAACTAGACTGTGCACTTGGATGTCACTGGAGTGTCACTATGGAGAGTTATAG	1732
Db	89964	AGAGCTGAACTAGACTGTGCACTTGGATGTCACTGGAGTGTCACTATGGAGAGTTATAG	89905
QY	1733	CAGCATCATACCAAGGCGCTTACTGTTCACATGAGGGCGCAAAACCAATTAACAGCCACTTC	1792
Db	89904	CAGCATCATACCAAGGCGCTTACTGTTCACATGAGGGCGCAAAACCAATTAACAGCCACTTC	89845
QY	1793	TTGGAAGAGGAATGCAAAAGCTTTGGGGGTGATGAGAAAGACCTTTTAAGATGATACCA	1852
Db	89844	TTGGAAGAGGAATGCAAAAGCTTTGGGGGTGATGAGAAAGACCTTTTAAGATGATACCA	89785
QY	1853	ATTAAACTGCGCTTGAAAGGGCACTAGGTGGG	1883
Db	89784	ATTAAACTGCGCTTGAAAGGGCACTAGGTGGG	89754
RESULT 10			
LOCUS	C0719642	981 bp	DNA
DEFINITION	Sequence 5576 from Patent WO02068579.	linear	PAT 03-FEB-2004
ACCESSION	C0719642		
VERSION	C0719642.1	GI:42280499	
KEYWORDS			

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 Venter, C. J., Adams, M. C., Li, P. W., and Myers, E. W. Kits, such as nucleic acid arrays, comprising a majority of humanecons or transcripts, for detecting expression and other uses thereof
JOURNAL	Patent: WO 02068579-A 5576 06-SEP-2002;
FEATURES	PE Corporation (NY) (US)
source	Location/Qualifiers 1..981 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	
Query Match	52.0%; Score 979.4; DB 6; Length 981;
Best Local Similarity	99.9%; Pred. No. 1.6e-160;
Matches	980; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	124 ATGGGCGCTATCCGGGTGACCCCGCCCGCGTCTGGAGAGAGATGATGCTTGACATAC 183
DB	1 ATGGCGCTATCCGGGTGACCCCGCCCGCGTCTGGAGAGAGATGATGCTTGACATAC 60
QY	184 TACGGGATGCTGTGCTTACCGTATGTTGAGGTGTGGCGCACTGACCGAGTGC 243
DB	61 TACGGGATGCTGTGCTTACCGTATGTTGAGGTGTGGCGCACTGACCGAGTGC 120
QY	244 GAGCTGAGAGCTCTTGCCCTTTCTGCTGGATGAGGCTCCTGGCCGCCCGGAGGCTTACC 303
DB	121 GAGCTGAGAGCTCTTGCCCTTTCTGCTGGATGAGGCTCCTGGCCGCCCGGAGGCTTACC 180
QY	304 CGGGCCCGCAGCGGCGCTGAGAGCTCCTGTGAGCTGAGACCGCGGCGGCAATGTCGAG 363
DB	181 CGGGCCCGCAGCGGCGCTGAGAGCTCCTGTGAGCTGAGACCGCGGCGGCAATGTCGAG 240
QY	364 AGCAACCTGCGAGCTGTGGGGCAATCTTCGCGGTGCGCCGCGCAAGCTGTGCGG 423
DB	241 AGCAACCTGCGAGCTGTGGGGCAATCTTCGCGGTGCGCCGCGCAAGCTGTGCGG 300
QY	424 CACCTGCGCGCGAAGCGGCGCGCGCAATGCTTCCAGAAAGCTATAGTATGCACTTCC 483
DB	301 CACCTGCGCGCGAAGCGGCGCGCGCAATGCTTCCAGAAAGCTATAGTATGCACTTCC 360
QY	484 AGCTCTTCAAGAAGGACAGAGGGTATAGCTGCGCGGTGCGGAGTCAAGAGTTCTGCA 543
DB	361 AGCTCTTCAAGAAGGACAGAGGGTATAGCTGCGCGGTGCGGAGTCAAGAGTTCTGCA 420
QY	544 AATTCTCAGCAGGGTCAGTGGGAGACAGGCTTCCCGCCCAACCAAGCGGACGGCGAGT 603
DB	421 AATTCTCAGCAGGGTCAGTGGGAGACAGGCTTCCCGCCCAACCAAGCGGACGGCGAGT 480
QY	604 CGGGGCGGGCCCAAGTGTGTGTGCCAGACGCGGCGGAGAGGGGCCCGACCGCACCCAG 663
DB	481 CGGGGCGGGCCCAAGTGTGTGTGCCAGACGCGGCGGAGAGGGGCCCGACCGCACCCAG 540
QY	664 CAGCAGTACAGGCCCGCGCAGACTTTCTCTAAGGCAAGTGAACCTTGACATCCGGCTC 723
DB	541 CAGCAGTACAGGCCCGCGCAGACTTTCTCTAAGGCAAGTGAACCTTGACATCCGGCTC 600
QY	724 CGGGTTCAGACAGATCTGCGAGCATGGGCGAGCTTGTGAGACAGAGGCGTGGCATCCGG 783
DB	601 CGGGTTCAGACAGATCTGCGAGCATGGGCGAGCTTGTGAGACAGAGGCGTGGCATCCGG 660
QY	784 CGGCCCCAGGCGCTGGCGCGGCACTGGAAGTGTGGGAGGCCACCGCAAGTCTGCGGC 843
DB	661 CGGCCCCAGGCGCTGGCGCGGCACTGGAAGTGTGGGAGGCCACCGCAAGTCTGCGGC 720
QY	844 TCAGAGGACCTGGGCTCTGTGTGTTGTGACATCAAGTTCTGAGAGCTTCTGATTTGGAC 903
DB	721 TCAGAGGACCTGGGCTCTGTGTGTTGTGACATCAAGTTCTGAGAGCTTCTGATTTGGAC 780

QY 904 GCCTTCTGGGCGACTAAGTGGCGCCCTGCTGAGAGCCCTGCGGGCGTGTCTTG 963
Db 781 GCCTTCTGGGCGACTAAGTGGCGCCCTGCTGAGAGCCCTGCGGGCGTGTCTTG 840
QY 964 ACTGAGGCGCTGCGAGAGGCTGTGGCGCGGAGGCTGTCCCTGCTGTGTGTGAT 1023
Db 841 ACTGAGGCGCTGCGAGAGGCTGTGGCGCGGAGGCTGTCCCTGCTGTGTGTGAT 900
QY 1024 GAGGCTGACTATGAGGCTGGCGCGGCGCTGTGTGATGAGAGAGAGGAGGCGG 1083
Db 901 GAGGCTGACTATGAGGCTGGCGCGGCGCTGTGTGATGAGAGAGAGGAGGCGG 960
QY 1084 CGCCCGACAGAGGCTCTCTGA 1104
Db 961 CGCCCGACAGAGGCTCTCTGA 981

RESULT 11
AF43591 981 bp mRNA linear PRI 25-FEB-2002
LOCUS Homo sapiens death effector domain-containing DNA-binding protein 2
DEFINITION (DED2) mRNA, complete cds.
ACCESSION AF43591.1 GI:17933266
VERSION AF43591.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 981)
AUTHORS Roch, W., Stenmer-Liwen, F., Pawlowski, K., Godzik, A. and Reed, J.C.
TITLE Identification and characterization of DED2, a death effector
domain-containing protein
J Biol. Chem. 277 (9), 7501-7508 (2002)
MEDLINE 21850646
PUBMED 11741985
REFERENCE 2 (bases 1 to 981)
AUTHORS Roch, W. and Reed, J.C.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) Burnham Institute, 10901 N. Torrey Pines
Rd, La Jolla, CA 92037, USA
FEATURES
source
1. .981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .981
/gene="DED2"
1. .981
/gene="DED2"
/codon_start=1
/product="death effector domain-containing DNA-binding
protein 2"
/protein_id="AA148220.1"
/db_xref="GI:17933267"
/translation="MALSGSTPAPCWEDECLDYGMLSLHMFVYVGGOLTECELT
LAFILDEAPAGAGIARBSGLLELEFRGOCDBENLLLOLARVLRHLLPLH
ARRRRRVSPEPYSYGTSSSKTBESCRRRSSSSANSQQQOMETGSPTRKRS
RGRPSGARRRRGAPAPQOSEPAPSSSEKAVTCIRLRVAEYCEHGPALQGV
SRPQALRQLDVFGQATAVLRSDLSVDCIKFSELSYLDAFWGDYLSGALLQAR
GVFLTEALREAVRERAVRLIVSYDEADYEAGRRRLIMEBEGRRPTENS"

ORIGIN
Query Match 52.0%; Score 979.4; DB 9; Length 981;
Best Local Similarity 99.9%; Pred. No. 1.6e-160;
Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 124 ATGGCGCTATCCGGGTGACCCCGCGCGCTGCTGGAGAGAGTGTGCTGGACTAC 183
Db 1 ATGGCGCTATCCGGGTGACCCCGCGCGCTGCTGGAGAGAGTGTGCTGGACTAC 60
QY 184 TACGGATGCTGTCTTCAACCGATGTTGAGGTGTGGCGGCAACTGACCGAGTGC 243

Db 61 TACGGATGCTGTCTTCAACCGATGTTGAGGTGTGGCGGCAACTGACCGAGTGC 120
QY 244 GAGCTGAGAGCTCCCGGCTTTCTGCTGGAGAGAGCTCTGCGCGCGGAGCTTGACC 303
Db 121 GAGCTGAGAGCTCCCGGCTTTCTGCTGGAGAGAGCTCTGCGCGCGGAGCTTGACC 180
QY 304 CGGCGCCGACGCGCTTGAAGAGCTCTGCTGAGAGAGAGCGCGGAGAGAGAGAG 363
Db 181 CGGCGCCGACGCGCTTGAAGAGCTCTGCTGAGAGAGAGCGCGGAGAGAGAGAG 240
QY 364 AGCAACTGCGGCTGCTGCGGCAACTCTGCGCGTGTGCGCGCGCAAGCTGCTCG 423
Db 241 AGCAACTGCGGCTGCTGCGGCAACTCTGCGCGTGTGCGCGCGCAAGCTGCTCG 300
QY 424 CACTGCGCGGCAAGCGGCGCGCGCGAGTGTCCAGAGAGCTATAGCACTTC 483
Db 301 CACTGCGCGGCAAGCGGCGCGCGCGAGTGTCTCCAGAGAGCTATAGCACTTC 360
QY 484 AGCTCTTCAAGAGAGAGAGAGAGTGTGCGCGTGTGCGCGAGTCAAGAGTTCGCA 543
Db 361 AGCTCTTCAAGAGAGAGAGAGAGTGTGCGCGTGTGCGCGAGTCAAGAGTTCGCA 420
QY 544 AATTCTCAGAGAGGTCAAGTGTGAGAGAGAGTGTCTCCCGCAAGAGAGAGAGT 603
Db 421 AATTCTCAGAGAGGTCAAGTGTGAGAGAGAGTGTCTCCCGCAAGAGAGAGAGT 480
QY 604 CGGCGCGCGCGCGAGTGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
Db 481 CGGCGCGCGCGCGAGTGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 664 CAGCAGTCAAGAGAGAGAGAGAGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 723
Db 541 CAGCAGTCAAGAGAGAGAGAGAGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 724 CGGCTGTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
Db 601 CGGCTGTGAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 784 CGGCGCGCGCGCGAGTGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
Db 661 CGGCGCGCGCGCGAGTGTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 844 TCAAGGAGAGCTGTGAGTGTGTGTGAGATGAGTGTCTGAGAGCTCTCTATCTGAG 903
Db 721 TCAAGGAGAGCTGTGAGTGTGTGTGAGATGAGTGTCTGAGAGCTCTCTATCTGAG 780
QY 904 GCCTTCTGGGCGACTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 963
Db 781 GCCTTCTGGGCGACTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840
QY 964 ACTGAGGCGCTGCGAGAGGCTGTGGCGCGGAGAGGCTGTGCTGTGTGTGTGAT 1023
Db 841 ACTGAGGCGCTGCGAGAGGCTGTGGCGCGGAGAGGCTGTGCTGTGTGTGTGAT 900
QY 1024 GAGGCTGACTATGAGGCTGGCGCGGCGCTGTGTGATGAGAGAGAGAGGAGGCGG 1083
Db 901 GAGGCTGACTATGAGGCTGGCGCGGCGCTGTGTGATGAGAGAGAGAGGAGGCGG 960
QY 1084 CGCCCGACAGAGGCTCTCTGA 1104
Db 961 CGCCCGACAGAGGCTCTCTGA 981

RESULT 12
AF457575 981 bp mRNA linear PRI 10-APR-2002
LOCUS Homo sapiens death effector domain-containing protein FLAME-3
DEFINITION (FLAME-3) mRNA, complete cds.
ACCESSION AF457575
VERSION AF457575.1 GI:20126793
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 981)
Zhan, Y., Hegde, R., Srinivasula, S.M., Fernandes-Alnemri, T. and
Alnemri, E.S.
Death effector domain-containing proteins DEDD and FLAME-3 form
nuclear complexes with the TRIF1C102 subunit of human transcription
factor I1IC
Cell Death Differ. 9 (4), 439-447 (2002)
JOURNAL 21965497
MEDLINE 11965497
PUBMED 2 (bases 1 to 981)
REFERENCE Alnemri, E.S.
AUTHORS Direct Submission
TITLE Submitted (11-DEC-2001) Microbiology and Immunology, Thomas
Jefferson University, 233 S. 10th Street, Philadelphia, PA 19107,
USA
FEATURES
source Location/Qualifiers
1. .981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
1. .981
/gene="FLAME-3"
1. .981
/gene="FLAME-3"
/codon_start=1
/product="death effector domain-containing protein
FLAME-3"
/protein_id="MAM10835.1"
/xref="GI:20126794"
/translation="MALSGSTPAPCWEBCDLDYGMILNMFVVGQLECELEL
LALDEAPGAGGGLARASGLLEFLERBCDDSNRLIGLLRYLARDHLPEHL
ARRRRPVSBERYSTSTSSSKRTBSGCRRRSSANSOQOMETGSPTRKORS
RGRPSGARRRRGAAPQOQSEPARPSSEKVTDLIRVAEYREHSPALQGV
SRPQALARQLDVGQATVLRSDIGSVVCDIKFSELSTLDLFMDYLSGALLQALR
GVFLTBALREAVRERLIVSDYEDYEXRRRLLMEEBGRPRPTES"

Query Match 51.6%; Score 971.4; DB 9; Length 981;
Best Local Similarity 99.4%; Pred. No. 4e-159;
Matches 975; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 124 ATGGCGCTATCGGGTGCAGCCCGGCGCGTGTGGGAGAGATGATGCTCGACTAC 183
DB 1 ATGGCGCTATCGGGTGCAGCCCGGCGCGTGTGGGAGAGATGATGCTCGACTAC 60
QY 184 TACGGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGCGCAACTGACCGAGTGC 243
DB 61 TACGGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGCGCAACTGACCGAGTGC 120
QY 244 GAGCTGAGGCTCTGGGCTTTTCTGCTGAGATGAGGCTCTTGGCGCGCGGAGGCTTACGCC 303
DB 121 GAGCTGAGGCTCTGGGCTTTTCTGCTGAGATGAGGCTCTTGGCGCGCGGAGGCTTACGCC 180
QY 304 CGGCGCGCGAGCGGCTAGAGCTCTGCTGAGAGTGTGGAGCGCGCGGAGAGTGTGGAG 363
DB 181 CGGCGCGCGAGCGGCTAGAGCTCTGCTGAGAGTGTGGAGCGCGCGGAGAGTGTGGAG 240
QY 364 AGCAACCTGCGGCTGTGGGCAACTCTTCTGCGGTGTGGCGCGCGCACTGCTGCCG 423
DB 241 AGCAACCTGCGGCTGTGGGCAACTCTTCTGCGGTGTGGCGCGCGCACTGCTGCCG 300
QY 424 CACCTGGCGCGCAAGCGCGCGCGCGCACTGCTTCCAGAAAGCTATGCTATGGCACTTCC 483
DB 301 CACCTGGCGCGCAAGCGCGCGCGCGCGCACTGCTTCCAGAAAGCTATGCTATGGCACTTCC 360
QY 484 AGCTCTTCAAGAAGAGCAGAGGGGTAGTGCCTGCGCGCGCGAGTCAACCACTTCCGA 543
DB 361 AGCTCTTCAAGAAGAGCAGAGGGGTAGTGCCTGCGCGCGCGAGTCAACCACTTCCGA 420

QY 544 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCTCCCTCCCAACCAAGCGGCGGAGT 603
DB 421 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCTCCCTCCCAACCAAGCGGCGGAGT 480
QY 604 CGGCGCGCGCGCGAGTGTGGTGCAGACGCGCGCGCGAGGGGCGCCAGCGCACCCGAG 663
DB 481 CGGCGCGCGCGCGAGTGTGGTGCAGACGCGCGCGCGAGGGGCGCCAGCGCACCCGAG 540
QY 664 CAGCAGTCAAGACCGCGCGCAGACCTTCTCTGAAGGCAAGTACCTGTGACATCCGGCTC 723
DB 541 CAGCAGTCAAGACCGCGCGCAGACCTTCTCTGAAGGCAAGTACCTGTGACATCCGGCTC 600
QY 724 CGGCTTCAAGCAGAGTACTGCGAGCATGGGCCAGCTTGGAGCAGAGGCGTGCATCCGG 783
DB 601 CGGCTTCAAGCAGAGTACTGCGAGCATGGGCCAGCTTGGAGCAGAGGCGTGCATCCGG 660
QY 784 CGGCGCGCGCGCGCGCGCGCGCGAGCTGAGCGTGTGGCAGAGCCAGCGCAGTGGCGGC 843
DB 661 CGGCGCGCGCGCGCGCGCGCGCGAGCTGAGCGTGTGGCAGAGCCAGCGCAGTGGCGGC 720
QY 844 TCAGAGGACCTGGGCTCTGTGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGAGC 903
DB 721 TCAGAGGACCTGGGCTCTGTGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGAGC 780
QY 904 GCCTTCTGGGCGAGTACTGAGTGGCGCGCTGTGCGAGCGCTTGGCGGCGCTTCTCTG 963
DB 781 GCCTTCTGGGCGAGTACTGAGTGGCGCGCGCTGTGCGAGCGCTTGGCGGCGCTTCTCTG 840
QY 964 ACTGAGGCGCGCGCGAGGCTGTGGCGCGCGAGGCGCTTGGCGGCGCTTGGCGGCGAT 1023
DB 841 ACTGAGGCGCGCGCGAGGCTGTGGCGCGCGAGGCGCTTGGCGGCGCTTGGCGGCGAT 900
QY 1024 GAGGCTGACTATGAGGCTGTGGCGCGCGCGCTGTGCTGATGAGAGGAGAGGCGGCGG 1083
DB 901 GAGGCTGACTATGAGGCTGTGGCGCGCGCGCGCTGTGCTGATGAGAGGAGAGGCGGCGG 960
QY 1084 CGCGCGCAGCAGAGGCTCTCTGA 1104
DB 961 CGCGCGCAGCAGAGGCTCTCTGA 981

RESULT 13
AX364852
LOCUS AX364852 1106 bp DNA linear PAT 15-FEB-2002
DEFINITION Sequence 3 from Patent WO0206315.
ACCESSION AX364852
VERSION AX364852.1 GI:18696741
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE Mintz, L., Freilich, S. and Bernstejn, J.
AUTHORS Novel nucleic acid and amino acid sequences
TITLE Patent: WO 0206315-A 3 24-JAN-2002;
JOURNAL Compugen Ltd. (IL)
FEATURES
source Location/Qualifiers
1. 1106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 50.3%; Score 946.4; DB 6; Length 1106;
Best Local Similarity 97.7%; Pred. No. 8.7e-155;
Matches 1036; Conservative 7; Mismatches 8; Indels 9; Gaps 8;

QY 27 AAAGATCCAAACCAATGCTGCGCGCGCTGCGCCAGAGATCAATCGACCCAGAACTCG 86
DB 10 AAAGATCCAAACCAATGCTGCGCGCGCTGCGCGCGAGATCAATCGACCCAGAACTCG 68

QY 87 CCGGGTTCTAGAGCTTTGTTCCGCTCCCTCCCGGGAAATGGCGCTATCCGGGTGACCCC 146
DB 69 CCGGGTTCTAGAGCTTTGTTCCGCTCCCTCCCGGGAAATGGCGCTATCCGGGTGACCCC 128
QY 147 GGGCCCGTGTGGGAGGAGATGATGCTGAGTACTAGCGGATGCTGCTTCAACG 206
DB 129 GGGCCCGTGTGGGAGGAGATGATGCTGAGTACTAGCGGATGCTGCTTCAACG 188
QY 207 TATGTTGAGAGTGTGGGCGGCAATGACGAGTGTGAGCTGAGCTTCTGCTTCT 266
DB 189 TATGTTGAGAGTGTGGGCGGCAATGACGAGTGTGAGCTGAGCTTCTGCTTCT 248
QY 267 GCTGATGAGCTCTTGGCGCGCGGAGGCTTACCCGGGCGCGACGGGCTTAAAGCT 326
DB 249 GCTGATGAGCTCTTGGCGCGCGGAGGCTTACCCGGGCGCGACGGGCTTAAAGCT 307
QY 327 CCGTGTGAGAGTGTGGGCGGCGGAGTGTGAGGAGAACTGTGGGCTGTGGGCA 386
DB 308 CCGTGTGAGAGTGTGGGCGGCGGAGTGTGAGGAGAACTGTGGGCTGTGGGCA 367
QY 387 ACTCTGCGGCTGTGGGCGGCGGAGCTGTGCGGCACTGTGGGCGGCAAGCGGCGCG 446
DB 368 ACTCTGCGGCTGTGGGCGGCGGAGCTGTGCGGCACTGTGGGCGGCAAGCGGCGCG 425
QY 447 GCCAGTGTCTCCAGAACGCTATAGCTATAGGCACTTCCAGCTCTTCAAGAGAGAGAGG 506
DB 426 GCCAGTGTCTCCAGAACGCTATAGCTATAGGCACTTCCAGCTCTTCAAGAGAGAGAGG 485
QY 507 TAGCTGCGGCTGTGGGCGGCGGAGTGTGAGGAGAACTGTGGGCTGTGGGCA 566
DB 486 TAGCTGCGGCTGTGGGCGGCGGAGTGTGAGGAGAACTGTGGGCTGTGGGCA 545
QY 567 GACAGGCTCCCGCCCAACCAAGCGGCGGAGTGTGGGCGGCGGCGGCGGCGGCGG 626
DB 546 GACAGGCTCCCGCCCAACCAAGCGGCGGAGTGTGGGCGGCGGCGGCGGCGGCGG 604
QY 627 CAGAGCGGCGGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 686
DB 605 CAGAGCGGCGGAGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
QY 687 TTCTCTGAGAGGCAAGTGTGATCTGTGATCTGTGCTGTGGGCTGTGGGCAAGTGTGCGA 746
DB 664 TTCTCTGAGAGGCAAGTGTGATCTGTGATCTGTGCTGTGGGCTGTGGGCAAGTGTGCGA 723
QY 747 GATGAGGCGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 806
DB 724 GATGAGGCGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 782
QY 807 GCTGAGCTGTGGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 866
DB 783 GCTGAGCTGTGGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 842
QY 867 TTGTGATCAATGATTTCTCAAGCTCTCTTATCTGGAAGCTTCTGTGGGCGGAGTGTGAG 926
DB 843 TTGTGATCAATGATTTCTCAAGCTCTCTTATCTGGAAGCTTCTGTGGGCGGAGTGTGAG 902
QY 927 TGGGCGGCTGTGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 986
DB 903 TGGGCGGCTGTGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 962
QY 987 GGGCGGAGAGGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1046
DB 963 GGGCGGAGGAGGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
QY 1047 GCGCGGCTGTGTGATGAG 1086
DB 1021 GCGCGGCTGTGTGATGAG 1060

RESULT 14
BD093311
LOCUS BD093311 909 bp DNA linear PAT 27-AUG-2002
DEFINITION Apoptosis Relating Factor.

ACCESSION BD093311
VERSION BD093311.1 GI:22638899
KEYWORDS MO 0104300-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Oca,T., Isogai,T., Nishikawa,T., Kawai,Y., Miyoshi,S. and Sato,S.
1 (bases 1 to 909)
Apoptosis Relating Factor
Patent: WO 0104300-A.1 18-JAN-2001;
JOURNAL HELIX RESEARCH INSTITUTE, TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI, SUSUMU SATO
COMMENT
PC C12N15/12, C07K14/47, C12N5/10, C12N1/21, C12N1/19, C12N1/15, C12P21/ PC
02, C07K16/18, C12P21/08, G01N33/53, G01N33/577
PC
CC
FH Key Location/Qualifiers
FT CDS (1)..(909).
FEATURES
source location/Qualifiers
1..909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 48.3%; Score 909; DB 6; Length 909;
Best Local Similarity 100.0%; Pred. No. 2,9e+148;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 124 ATGGCGCTATCCGGGCTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183
DB 1 ATGGCGCTATCCGGGCTGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 184 TACGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGGCGGCGGCAATGACGAGTGC 243
DB 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGGCGGCGGCAATGACGAGTGC 120
QY 244 GAGCTGAGCTCTTGCTGCTTCTGTGATGAGGCTCTGAGCGCGCGGAGGCTTAGCC 303
DB 121 GAGCTGAGCTCTTGCTGCTTCTGTGATGAGGCTCTGAGCGCGCGGAGGCTTAGCC 180
QY 304 CGGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 363
DB 181 CGGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 364 AGCAACTGTGGGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 423
DB 241 AGCAACTGTGGGCTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 424 CACCTGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
DB 301 CACCTGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 484 AGCTTTCAAGAGGAG 543
DB 361 AGCTTTCAAGAGGAG 420
QY 544 AATTCTCAGAGAGGATGAGTGGAGACAGAGCTCCCGCCCAACCAAGCGGCGGCGGAGT 603
DB 421 AATTCTCAGAGAGGATGAGTGGAGACAGAGCTCCCGCCCAACCAAGCGGCGGCGGAGT 480
QY 604 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663

Db 481 CGGGGCGGCGCCAGTGTGTGTCACAGCGGCGGAGAGGGGCCCGCCAGCCGACCCAG 540
QY 664 CAGCAGTCAGAGCCGCCAGACCTTCTCTGAAAGCAAGTGAATCTGTGACATCCGGCTC 723
Db 541 CAGCAGTCAGAGCCGCCAGACCTTCTCTGAAAGCAAGTGAATCTGTGACATCCGGCTC 600
QY 724 CGGGTTCGAGCAGAGTACCTGCGAGCATGGGCGCAGCTTGGAGAGAGGCGGTGGATCCCG 783
Db 601 CGGGTTCGAGCAGAGTACCTGCGAGCATGGGCGCAGCTTGGAGAGAGGCGGTGGATCCCG 660
QY 784 CGGCCCCAGGCGCTGGCGCGCGAGCTGAGAGTGTGGGAGGCGCACCGCAGTGCAGC 843
Db 661 CGGCCCCAGGCGCTGGCGCGCGAGCTGAGAGTGTGGGAGGCGCACCGCAGTGCAGC 720
QY 844 TCAGAGGAGCTGGGCTCTGTGTGTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGC 903
Db 721 TCAGAGGAGCTGGGCTCTGTGTGTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGC 780
QY 904 GCTTCTGGGGGCACTACCTGAGTGGCGCCTTCTGTGAGGCGCTGGGGGCGGTTCCTG 963
Db 781 GCTTCTGGGGGCACTACCTGAGTGGCGCCTTCTGTGAGGCGCTGGGGGCGGTTCCTG 840
QY 964 ACTGAGGCGCTGGGAGGAGCTGTGGGCGGAGAGGAGTGTGCTGCTGTGCTGAGTGTGAT 1023
Db 841 ACTGAGGCGCTGGGAGGAGCTGTGGGCGGAGAGGAGTGTGCTGCTGTGCTGAGTGTGAT 900
QY 1024 GAGGCTGAC 1032
Db 901 GAGGCTGAC 909

RESULT 15
AX364851 1084 bp DNA linear PAT 15-FEB-2002
LOCUS AX364851
DEFINITION Sequence 2 from Patent WO0206315.
ACCESSION AX364851
VERSION AX364851.1 GI:18696740
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukayocra; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Mintz, L., Freilich, S. and Bernstein, J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 2 24-JAN-2002;
CompuGen Ltd. (IL)
FEATURES
Source Location/Qualifiers
1..1084
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 47.3%; Score 902.4; DB 6; Length 1084;
Best Local Similarity 97.2%; Pred. No. 4e-147;
Matches 984; Conservative 7; Mismatches 13; Indels 8; Gaps 7;

QY 75 GCCAGAAATGCGCGGCTTCTGAGCTTTCGCCCTCCCGCGGAAATGGCGCTATC 134
Db 35 GGCACCTTTGGCCGGGTTCTGAGCTTGTCCGCTCCCGCGGAAATGGCGCTATC 94
QY 135 CGGGTGCAGCCCGGCGCGCTGCTGGAGAGAGATGAGTGCCTGGAATCACTACGGGATGCT 194
Db 95 CGGGTGCAGCCCGGCGCGCTGCTGGAGAGAGATGAGTGCCTGGAATCACTACGGGATGCT 154
QY 195 GTGCTTACCGATGTTGAGGTGTGGCGGCGCAACTGACCGAGTGCAGCTGAGCT 254
Db 155 GTGCTTACCGATGTTGAGGTGTGGCGGCGCAACTGACCGAGTGCAGCTGAGCT 214
QY 255 CTTGGCTTTCTGTGTGATGAGGCTCTGGCGCGCGCGGAGGCTTAAGCCCGGCGCGAG 314
Db 215 CTTGGCTTTCTGTGTGATGAGGCTCTGGCGCGCGCGGAGG-TTAAAGCCGGCGCGAG 273

QY 315 CGGCTTAAGCTTCTGTGTGAGTGTGAGCGCGCGGCGAGTGGCGGAGAGCAACTGCG 374
Db 274 CGGCTTAAGCTTCTGTGTGAGTGTGAGCGCGCGGCGAGTGGCGGAGAGCAACTGCG 333
QY 375 GCTGTGGGGCAACTCTGCGGTGTGGCGCGGCGAGCACTGTGCGCAGCTGGCGCG 434
Db 334 GCTGTGGGGCAACTCTGCGGTGTGGCGCGGCGAGCACTGTGCGCAGCTGGCGCG 393
QY 435 CAAGCGCGCGCGGCGAGTGTCTCAAGACGTATAGTATGACACTCCAGCTTTCAAA 494
Db 394 CAA--AGCGCGGCGAGTGTCTCAAGACGTATAGTATGACACTCCAGCTTTCAAA 451
QY 495 GAGAGACAGGGTGTGCTGCGGTGTGCGGTGTGGAGTCAAGCAATTTCTGCAAA 554
Db 452 GAGAGACAGGGTGTGCTGCGGTGTGCGGTGTGGAGTCAAGCAATTTCTGCAAA 511
QY 555 GGGTCAAGGGAGACAGGCTCCCGCCCAACCAAGCGGCGAGCGGAGTCCGGGCGCGCC 614
Db 512 GGGTCAAGGGAGACAGGCTCCCGCCCAACCAAGCGGCGAGCGGAGTCCGGGCGCGCC 570
QY 615 CAGTGTGTGTGCGAGACGCGCGGCGAGAGAGGGCCCCAGCGCACTCCAGCAGTCAAG 674
Db 571 CAGTGTGTGTGCGAGACGCGCGGCGAGAGAGGGCCCCAGCGCACTCCAGCAGTCAAG 629
QY 675 GCGCGCCAGACCTTCTCTGAAAGGCAAGTGAATCTGTGACATCCGGCTCCGGGTTGAGC 734
Db 630 GCGCGCCAGACCTTCTCTGAAAGGCAAGTGAATCTGTGACATCCGGCTCCGGGTTGAGC 689
QY 735 AAGATCTGCGAGCAATGGGCGCAGCCTTGGAGAGAGGGCGTGGATCCCGCGGCGCGAGC 794
Db 690 AAGATCTGCGAGCAATGGGCGCAGCCTTGGAGAGAGGGCGTGGATCCCGCGGCGAGC 748
QY 795 GCTGGCGCGGCGAGCTGAGCGTGTGTGGGAGGCGCACCGCAGTGTGCTCAAGGAGCT 854
Db 749 GCTGGCGCGGCGAGCGTGTGTGGGAGGCGCACCGCAGTGTGCTCAAGGAGCT 808
QY 855 GGGCTGTGTGTGTGTGACATCAAGTGTCAAGGCTCTCTATCTGAGAGCTTCTGGGG 914
Db 809 GGGCTGTGTGTGTGTGACATCAAGTGTCTCAAGGCTCTCTATCTGAGAGCTTCTGGGG 868
QY 915 CGACTACCTGAGTGGCGCGCTGCTGTGAGGCGCTGGCGGCGGTCTGACTGAGGCGCT 974
Db 869 CGACTACCTGAGTGGCGCGCTGCTGTGAGGCGCTGGCGGCGGTCTGACTGAGGCGCT 928
QY 975 GCGAGAGGCTGTGGCGCGGAGGCTGTGCGCTGTGTGCTGAGTGTGAGGCTGACTTA 1034
Db 929 GCGAGAGGCTGTGGCGCGGAGG-CTGTGTGCGCTGTGTGAGTGTGAGGCTGACTTA 987
QY 1035 TGAAGCTGGCGCGCGCGCTTGTGCTGATGAGAGAGAGAGGCGGCGCGC 1086
Db 988 TGAAGCTGGCGCGCGCGCTTGTGCTGATGAGAGAGAGAGAGGCGGCGCGC 1038

Search completed: March 23, 2005, 13:22:43
Job time : 8487.2 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 01:57:52 ; Search time 32.3945 Seconds
(without alignments)
5399.743 Million cell updates/sec

Title: US-10-030-271-1
Perfect score: 303
Sequence: 1 atggcgctaccgggtcgac.....tcagtcgtgacgagcgctgac 909

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Xgapop 60.0 , Xgapext 60.0
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565918

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-Q=/cgnr2/USPRO_pool/p/US10030271/runatc_21032005_153547_26829/app_query.fasta_1.3150
-DB=PIR_79 -QFMT=ascan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030271 @CGN 1 1 151 @runat 21032005 153547 26829 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	3.3	235	2 T30656	hypothetical prote
2	10	3.3	1460	1 EDBE1F	immediate-early pr
3	127	9	3.0	2 A72522	hypothetical prote
4	375	9	3.0	2 C75584	hypothetical prote
5	370	9	3.0	1 TVHUT1	transforming prote
6	454	9	3.0	2 AE1970	hypothetical prote
7	75	8	2.6	2 AE2664	ATP synthase C cha
8	75	8	2.6	2 D97446	ATP synthase C cha
9	134	8	2.7	2 T44984	methylinomy1-COA
10	158	8	2.7	2 T26444	hypothetical prote
11	170	8	2.6	2 C84492	hypothetical prote
12	170	8	2.6	2 P83325	hypothetical prote
13	195	8	2.6	2 E75462	hypothetical prote
14	195	8	2.6	2 S45627	acidic endoprotein

15	8	2.6	197	2 C70601	hypothetical prote
16	8	2.7	206	2 AG0466	probable homoserin
17	8	2.7	213	2 E75525	transcription regu
18	8	2.6	219	2 A57694	transcription fact
19	8	2.7	219	2 C87371	hypothetical prote
20	8	2.6	224	2 B87197	hypothetical prote
21	8	2.6	224	2 E70790	probable transcript
22	8	2.6	236	1 T46967	diheme cytochrome
23	8	2.6	262	2 C70648	probable nuot prot
24	8	2.7	265	2 AH0775	hydroxyethylthiaz
25	8	2.6	266	2 AE0156	probable transcrip
26	8	2.6	279	2 B70328	hypothetical prote
27	8	2.7	279	2 T34848	hypothetical prote
28	8	2.6	282	2 S46793	vacuolar protein s
29	8	2.7	288	2 P84823	homeotic protein e
30	8	2.7	314	2 C87667	ABC transporter, A
31	8	2.6	339	2 S20880	homeotic protein H
32	8	2.6	342	2 S18649	homeotic protein H
33	8	2.6	363	2 T34931	hypothetical prote
34	8	2.7	366	2 T51339	mitogen-activated
35	8	2.7	375	2 T05707	phosphate transpor
36	8	2.6	376	2 C75580	adenine deaminase
37	8	2.7	381	2 S08296	beta-lactamase (EC
38	8	2.7	381	2 S45109	beta-lactamase (EC
39	8	2.6	381	2 AD0897	conserved hypothet
40	8	2.7	394	2 E87675	hypothetical prote
41	8	2.6	397	2 T08345	hypothetical prote
42	8	2.7	400	1 SYZMW1	maringenin-chalcon
43	8	2.6	402	2 A45056	osteogenic protein
44	8	2.6	417	2 E86858	cell division prot
45	8	2.7	435	2 AG3257	CBS domain contain
46	8	2.6	460	2 T07868	DNA binding protei
47	8	2.7	463	2 S77558	hypothetical prote
48	8	2.6	474	2 D64485	TRK system potassi
49	8	2.6	475	2 A38340	66k glycoprotein p
50	8	2.6	477	1 ORH0B1	beta-1-adrenergic
51	8	2.6	480	2 I53053	beta 1 adrenergic
52	8	2.6	513	1 S43342	flavonoid 3',5'-hy
53	8	2.7	521	2 F87317	hypothetical prote
54	8	2.6	528	2 AC3236	hypothetical prote
55	8	2.7	598	2 S65770	maltooligosyltraha
56	8	2.7	738	1 TETHUM	melanotransferin
57	8	2.6	761	2 S20458	pqqf protein - K1e
58	8	2.7	767	2 A35645	major surface prot
59	8	2.6	775	2 I49759	hepatocyte growth
60	8	2.6	786	2 S22155	oncogene 1 (tre-2
61	8	2.6	798	1 B44051	infected cell prot
62	8	2.6	828	2 T33481	hypothetical prote
63	8	2.6	862	2 T46289	hypothetical prote
64	8	2.7	877	2 T35861	probable large sec
65	8	2.6	914	2 T00757	probable ubiquitin
66	8	2.6	928	2 T04192	hypothetical prote
67	8	2.6	937	2 T04194	hypothetical prote
68	8	2.6	963	2 T09478	ubiquitin thiolest
69	8	2.7	1073	2 T38763	hypothetical prote
70	8	2.6	1089	2 S22158	transforming prote
71	8	2.6	1095	2 B83471	probable pyruvate
72	8	2.7	1162	2 T51040	hypothetical prote
73	8	2.7	1222	2 G59100	hypothetical prote
74	8	2.7	1293	2 T30871	orsellinic acid sy
75	8	2.7	1704	2 T43141	vitellinogenin 1 - m
76	8	2.6	1733	1 B45344	probable nuclear a
77	8	2.7	1788	2 T31095	vitellinogenin precu
78	8	2.6	1958	2 B40505	hypothetical prote
79	8	2.7	2279	2 T42531	acetyl-CoA carboxy
80	8	2.7	2280	2 T38906	acetyl-CoA carboxy
81	8	2.7	4848	2 T30289	pristinamycin I sy
82	8	2.3	12	2 A60528	insulin-like growt
83	8	2.3	12	2 A33520	inhibitory diffusi
84	7	2.3	20	2 S38288	50k allergen - per
85	7	2.3	20	2 A61505	pollen allergen Ph
86	7	2.3	34	2 S02382	probable membrane
87	7	2.3	47	2 T29970	hypothetical prote

US-10-030-271-1 (1-909) x F83325 (1-170)
QY 777 GGACGCTTGTGGGCGACTACT 800
DB 130 GYATGLeuLeuGlyATGLeuPro 137

RESULT 13
E75462
hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: E75462
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.V.;
M.; Shen, M.; Vamathevan, J.U.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75462

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-195 <WHI>

A:Cross-references: UNIPROT:Q9RYX4; GB:AE001942; GB:AE000513; NID:g6458611; PIDN:AAF1047

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0897

A:Map position: 1

C:Superfamily: Deinococcus radiodurans hypothetical protein DR0897

Alignment Scores:

Pred. No.: 103 Length: 195
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x E75462 (1-195)

QY 99 GGGCGGCGACTGACCGGAGTGCGA 122

DB 60 GYATGLeuLeuATGATGATG 67

RESULT 14

S45627
acidic endoprotease precursor - Myxococcus xanthus (strain DK101)

C:Species: Myxococcus xanthus

A:Variety: strain DK101

C:Date: 10-Dec-1994 #sequence_revision 26-Apr-1996 #text_change 09-Jul-2004

C:Accession: S45627; S62857

R:Lucas, N.; Mazaud-Aujard, C.; Bremaud, L.; Genetempo, Y.; Julien, R.

Eur. J. Biochem. 222, 247-254, 1994

A:Title: Protein purification, gene cloning and sequencing of an acidic endoprotease frc

A:Reference number: S45627; MUID:94291618; PMID:8020464

A:Accession: S45627

A:Molecule type: DNA

A:Residues: 1-195 <LUC>

A:Cross-references: UNIPROT:Q50902; EMBL:X75892; NID:g516391; PIDN:CA53499.1; PID:g5163

A:Experimental source: strain DK101

A:Accession: S62857

A:Molecule type: protein

A:Residues: 65-101 <LUM>

C:Genetics:

A:Gene: Maep

F1-29/Domain: signal sequence #status predicted <SIG>

F30-64/Domain: propeptide #status predicted <PRO>

F65-195/Product: acidic endoprotease #status experimental <MAT>

Alignment Scores:

Pred. No.: 103 Length: 195
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x S45627 (1-195)

QY 282 CCGCAGCGACTGCTGCCGACCT 305

DB 150 ProProATProAlaAlaAlaPro 157

RESULT 15

C70601
hypothetical protein RV0992c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: C70601

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.B.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: C70601

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-197 <COL>

A:Cross-references: UNIPROT:O05575; GB:Z94752; GB:AL123456; NID:g3261731; PIDN:CAB08152.

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0992c

C:Superfamily: human 5-formyltetrahydrofolate cyclo-ligase

Alignment Scores:

Pred. No.: 103 Length: 197
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x C70601 (1-197)

QY 166 GCCGAGGCTTACCGGCGCCGCG 189

DB 95 AlaGlyGlyLeuAlaAlaAlaAla 102

RESULT 16

AG0466

probable homoserine/homoserine lactone efflux protein rhbB [imported] - Yersinia pestis

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AG0466

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tiltball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AG0466

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-206 <KUR>

A:Cross-references: UNIPROT:Q8ZAH0; GB:AL590842; PIDN:CAC93299.1; PID:g15981746; GSPDB:G

C:Genetics:

A:Gene: rhbB

Alignment Scores:

Pred. No.: 102 Length: 206
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

A;Molecule type: DNA
A;Residues: 1-265 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02526.1; PID:G16503387; GSPDB:GN00176
C;Genetics:
A;Gene: STY2376
C;Superfamily: hydroxyethylthiazole kinase; hydroxyethylthiazole kinase homology
C;Keywords: phosphotransferase

Alignment Scores:
Pred. No.: 96.1 Length: 265
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x AH0775 (1-265)

QY 243 GCTTCGCCGACATGCCCCGCGCG 220
Db 150 AlaleuAlaAlaLeuProAlaAla 157

RESULT 25
AE0156
probable transcription regulatory protein YP01279 [imported] - Yersinia pestis (strain C
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AE0156
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Hl, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0156
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-266 <KOR>
A;Cross-references: UNIPROT:Q8ZGL0; GB:AL590842; PIDN:CAC90112.1; PID:G15979332; GSPDB:C
C;Genetics:
A;Gene: YP01279
C;Superfamily: regulatory protein uxuR 2

Alignment Scores:
Pred. No.: 96 Length: 266
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x AE0156 (1-266)

QY 435 TCAGTGGAGACAGGCTCCCCCCC 458
Db 40 SerValGlyAparGleuProPro 47

RESULT 26
B70328
hypothetical protein aq_313 - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: B70328
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oy
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: B70328
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-279 <AGP>
A;Cross-references: UNIPROT:O66654; GB:AE000683; NID:G2982996; PIDN:AAC06617.1; PID:G298

A;Experimental source: strain VFS
C;Genetics:
A;Gene: aq_313
C;Superfamily: hypothetical protein HP0152

Alignment Scores:
Pred. No.: 95 Length: 279
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x B70328 (1-279)

QY 865 GCGCGGAGGCTGTGCTGCTG 888
Db 248 GlyArgGlnAlaValArgLeu 255

RESULT 27
T34848
probable transcription regulator - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34848
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A;Reference number: Z21559
A;Accession: T34848
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-279 <OLI>
A;Cross-references: UNIPROT:Q9Z5A8; EMBL:AL035478; PIDN:CAB36602.1; GSPDB:GN00070; SCOEDF
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC2G5.15C

Alignment Scores:
Pred. No.: 95 Length: 279
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x T34848 (1-279)

QY 679 GCGCGAGCGCTGGCGCGCG 656
Db 162 AlaProAlaProGlyAlaAlaGly 169

RESULT 28
S46793
vacuolar protein sorting protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein PEP1; protein VPS29; protein YHR012w
C;Species: Saccharomyces cerevisiae
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: S46793
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid l2825.
A;Reference number: S46774
A;Accession: S46793
A;Molecule type: DNA
A;Residues: 1-282 <DUZ>
A;Cross-references: UNIPROT:P38759; EMBL:U10400; NID:G500701; PID:G500712; GSPDB:GN00008
C;Genetics:
A;Gene: VPS29; MIPS:YHR012w
A;Cross-references: MIPS:YHR012w; SGD:S0001054
A;Map position: 8R
A;Intons: 16/3
C;Superfamily: VPS29-like phosphoesterase-related protein

Alignment Scores:

Pred. No.:	94.8	Length:	282
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.64%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-1 (1-909) x S46793 (1-282)

Qy 670 GCGCTGGCGGCGAGCTGAGCTG 693

Db 120 AAlaAlaAlaArgGlnLeuAspVal 127

RESULT 29

F48423

homeotic protein engrailed 2 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Aug-2004

C:Accession: F48423

R:Logan, C.; Hanks, M.C.; Noble-Topham, S.; Nallainathan, D.; Provart, N.J.; Joyner, A.L

Dev. Genet. 13, 345-358, 1992

A:Title: Cloning and sequence comparison of the mouse, human, and chicken engrailed gene

A:Reference number: A48423; MUID:93185339; PMID:1363401

A:Accession: F48423

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

C:Superfamily: homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:200-256/Domain: homeobox homology <HOX>

Alignment Scores:

Pred. No.:	94.3	Length:	288
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.67%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-1 (1-909) x F48423 (1-288)

Qy 685 GCTGCGGCGGCGAGCTGAGCTG 662

Db 87 AAlaAlaAlaProAlaProGlyAla 94

RESULT 30

C87667

ABC transporter, ATP-binding protein CC3373 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 16-Aug-2004

C:Accession: C87667

R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolot

n, J.; Birmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87667

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-314 <STO>

A:Cross-references: UNIPROT:Q9A333; GB:AE005673; NID:g13425079; PIDN:AAK25335.1; GSPDB:C

C:Gene: CC3373

C:Superfamily: ATP-binding cassette homology

Alignment Scores:

Pred. No.:	92.5	Length:	314
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.67%	Indels:	0

DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x C87667 (1-314)

Qy 240 CTCGCGGCGAGCTGCGGCGGCTG 217

Db 244 LeuAlaAlaLeuProAlaAlaLeu 251

RESULT 31

S20880

homeotic protein Hox 4.5 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Aug-2004

C:Accession: S20880; S09569; S09398

R:Renucci, A.; Zappavigna, V.; Zakany, J.; Izpisua-Belmonte, J.C.; Buerki, K.; Duboule,

EMBO J. 11, 1459-1468, 1992

A:Title: Comparison of mouse and human HOX-4 complexes defines conserved sequences invol

A:Reference number: S20879; MUID:92224884; PMID:1348690

A:Accession: S20880

A:Molecule type: DNA

A:Residues: 1-339 <REN>

A:Cross-references: UNIPROT:P28357; EMBL:X62669; NID:g51414; PIDN:CAA44542.1; PID:g51416

R:Duboule, D.; Dolle, P.

EMBO J. 8, 1497-1505, 1989

A:Title: The structural and functional organization of the murine HOX gene family resemb

A:Reference number: S09569; MUID:89356621; PMID:2569969

A:Accession: S09569

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 272-331 <DUB>

A:Cross-references: EMBL:X14714; NID:g51427; PIDN:CAB57813.1; PID:g6015583

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1989

R:Dolle, P.; Duboule, D.

EMBO J. 8, 1507-1515, 1989

A:Title: Two gene members of the murine HOX-5 complex show regional and cell-type specif

A:Reference number: S09398; MUID:89356622; PMID:2569970

A:Accession: S09398

A:Molecule type: DNA

A:Residues: 272-331 <DOL>

A:Cross-references: GB:X14714; GB:M21040; NID:g51427; PIDN:CAB57813.1; PID:g6015583

C:Genetics:

A:Gene: Hox-4.5

A:Introns: 260/1

C:Superfamily: homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:273-329/Domain: homeobox homology <HOX>

Alignment Scores:

Pred. No.:	91	Length:	339
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.64%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-1 (1-909) x S20880 (1-339)

Qy 358 TCCAGCTTTCAGAGAGAGAGAG 381

Db 173 SerSerSerSerIysArgThrGlu 180

RESULT 32

S18649

homeotic protein HOX D9 - human

N:Alternate names: homeotic protein Hox 4C; homeotic protein Hox 5.2

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Aug-2004

C:Accession: S18649; S09569; S14935; A32830

R:Zappavigna, V.; Renucci, A.; Izpisua-Belmonte, J.C.; Uribe, G.; Peschle, C.; Duboule,

EMBO J. 10, 4177-4187, 1991

A:Title: HOX4 genes encode transcription factors with potential auto- and cross-regulato

A:Reference number: S18649; MUID:92097538; PMID:1756725

A:Accession: S18649

A;Molecule type: mRNA
A;Residues: 1-342 <ZAP>
A;Cross-references: UNIPROT:P28356; EMBL:X59372; NID:G32390; PIDN:CAA42016.1; PID:G32391
N;Alternate names: MAP kinase kinase 4
R;Oliver, G.; Sidell, N.; Fiske, W.; Heinzmann, C.; Mohandas, T.; Sparkes, R.S.; De Robe
Genes Dev. 3, 641-650, 1989
A;Title: Complementary homeo protein gradients in developing limb buds.
A;Reference number: A32830; MUID:89306602; PMID:2568311
A;Accession: S05958
A;Molecule type: DNA
A;Residues: 264-265 'A', 267-342 <OLI>
A;Cross-references: EMBL:X15506; NID:G32397; PIDN:CAA33528.1; PID:G32398
R;Acampora, D.; d'Esposito, M.; Patelletti, A.; Pannese, M.; Migliaccio, E.; Morelli, F.; S
Nucleic Acids Res. 17, 10385-10402, 1989
A;Title: The human HOX gene family.
A;Reference number: S07541; MUID:90098876; PMID:2574852
A;Accession: S14935
A;Molecule type: DNA
A;Residues: 275-340 <ACA>
C;Genetics:
A;Gene: GDB:HOXD9
A;Cross-references: GDB:120678; OMIM:142982
A;Map position: 2q31-2q31
A;Intons: 263/1
C;Superfamily: homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;276-332/homein: homeobox homology <HOX>

Alignment Scores:
Pred. No.: 90.8 Length: 342
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x S18649 (1-342)

QY 358 TCACGCTCTTCAAGAGACAGAG 381
|||||
Db 175 SerserserlyrargThrGlu 182

RESULT 33

T34931
hypothetical protein SC3F9.09 SC3F9.09 - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Raebstream, M.A.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21562
A;Accession: T34931
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-363 <SER>
A;Cross-references: UNIPROT:O69949; EMBL:AL023862; PIDN:CAA19632.1; GSPDB:GN00070; SCOEI
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEI:SC3F9.09

Alignment Scores:
Pred. No.: 89.6 Length: 363
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x T34931 (1-363)

QY 7 CTATCCGGGTGACCCGCGCCCG 30
|||||
Db 269 leuserglyserThrProAlapro 276

RESULT 34
T51339
mitogen-activated protein kinase (EC 2.7.1.-) 4 [validated] - Arabidopsis thaliana
N;Alternate names: MAP kinase kinase 4
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
A;Accession: T51339
R;Ichimura, K.; Mizoguchi, T.; Hayashida, N.; Seki, M.; Shinozaki, K.
DNA Res. 5, 341-348, 1998
A;Title: Molecular cloning and characterization of three cDNAs encoding putative mitogen-
A;Reference number: Z25272; MUID:99156228; PMID:10048483
A;Accession: T51339
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-366 <ICH>
A;Cross-references: UNIPROT:O80397; EMBL:AB015315; PIDN:BA28830.1
A;Experimental source: strain Columbia
C;Genetics:
A;Gene: ATMKK4
C;Function:
A;Description: (EC 2.7.1.-); mitogen-activated protein kinase kinase [validated, MUID:99]
C;Superfamily: kinase-related transforming protein; protein kinase homology
C;Keywords: phosphotransferase; protein kinase

Alignment Scores:
Pred. No.: 89.5 Length: 366
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x T51339 (1-366)

QY 796 AGTCGCCCGAGAGGTCACAGAT 773
|||||
Db 16 SerrargProalrargargProasp 23

RESULT 35

T05707
phosphate transport protein G7, mitochondrial - soybean
C;Species: Glycine max (soybean)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
R;Takabatake, R.; Hata, S.; Taniguchi, M.; Kouchi, H.; Sugiyama, T.; Izui, K.
submitted to the EMBL Data Library, July 1998
A;Description: Isolation and characterization of cDNAs encoding mitochondrial phosphate t
A;Reference number: Z15423
A;Accession: T05707
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-375 <TAK>
A;Cross-references: UNIPROT:O80412; EMBL:AB016063; PIDN:BA31582.1
A;Experimental source: cultivar Akisengoku; root tip
C;Genetics:
A;Gene: nuclear
A;Superfamily: ADP, ATP carrier protein; ADP, ATP carrier protein repeat homology
C;Keywords: mitochondrion; phosphate transport

Alignment Scores:
Pred. No.: 89 Length: 375
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x T05707 (1-375)

QY 688 CCAGCTGCGCGCAGCGCCTGGG 665
|||||
Db 362 ProAlaAlaProAlaProGly 369

A/Accession: AD0897
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-381 <P&R>
A/Cross-references: GB:AL513382; PIDN:CAD07770.1; PID:G16504319; GSPDB:GN00176
C/Genetic8
A/Gene: STY3429
C/Superfamily: yhad protein

Alignment Scores:
Pred. No.: 88.7 Length: 381
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x AD0897 (1-381)

QY 660 GCGGCCCCGAGCGCTGGCGCGCA 683
Db 243 AlAlAlProGlyAlaGlyAlaAla 250

RESULT 40
E87675
hypothetical protein CC3439 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: E87675
R/Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.O.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: E87675
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-394 <STO>
A/Cross-references: UNIPROT:Q9A2W7; GB:AE005673; NID:G13425155; PIDN:AAK25401.1; GSPDB:C
C/Genetics:
A/Gene: CC3439

Alignment Scores:
Pred. No.: 88 Length: 394
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x E87675 (1-394)

QY 685 GCTGCGGCGAGCGCTGGCGGC 662
Db 13 AlAlAlAlProAlaProGlyAla 20

RESULT 41
T08345
hypothetical protein H1502 [imported] - Halobacterium sp. (strain NRC-1) plasmid pNRC100
C/Species: Halobacterium sp.
A/Variety: strain NRC-1
C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C/Accession: T08345
R/Ng, W.V.; Clufo, S.A.; Smith, T.M.; Bungarner, R.E.; Baekin, D.; Faust, J.; Hall, B.;
Genome Res. 8, 1131-1141, 1998
A/Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or m
A/Reference number: 216408; MUID:99063795; PMID:9847077
A/Accession: T08345
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-397 <DAS>
A/Cross-references: UNIPROT:O52032; EMBL:AF016485; NID:G2822278; PID:G2822406; HALOSP:HL

A/Experimental source: strain NRC-1
C/Genetics:
A/Gene: HALOSP:H1502
A/Genome: plasmid pNRC100
C/Superfamily: Halobacterium plasmid pNRC100 hypothetical protein H1502

Alignment Scores:
Pred. No.: 87.9 Length: 397
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x T08345 (1-397)

QY 630 GCCGCGCTTGAGCAGCGCTGGC 653
Db 85 AlAserLeuGlyAlaGlyArgGly 92

RESULT 42
SYZMW1
naringenin-chalcone synthase (EC 2.3.1.74) whp1 - maize
C/Species: Zea mays (maize)
C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C/Accession: S16599
R/Franken, P.; Niesbach-Kloessen, U.; Weydemann, U.; Marschal-Drouard, L.; Seadler, H.; v
EMBO J. 10, 2605-2612, 1991
A/Title: The duplicated chalcone synthase genes C2 and Whp (white pollen) of Zea mays are
gene in.
A/Reference number: S16599; MUID:91330885; PMID:1714383
A/Accession: S16599
A/Molecule type: DNA
A/Residues: 1-400 <FRA>
A/Cross-references: UNIPROT:P24824; EMBL:X60204; NID:G22511; PIDN:CAA42763.1; PID:G22512
C/Genetics:
A/Gene: whp1
A/Map position: 2L
A/Intons: 64/1
C/Superfamily: Type III polyketide synthase
C/Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Alignment Scores:
Pred. No.: 87.8 Length: 400
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 1 Gaps: 0

US-10-030-271-1 (1-909) x SYZMW1 (1-400)

QY 304 GGTGCGGAGCAGTCTGGCGG 281
Db 221 G1yAlAlAlaGlyArgGlyGly 228

RESULT 43
A45056
osteogenic protein 2 precursor - human
N/Alternate names: bone morphogenetic protein 8; OP-2
C/Species: Homo sapiens (man)
C/Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A45056
R/Oerkavna, E.; Schlegelberg, P.N.; Jin, D.F.; Clifford, G.M.; Warren, F.D.; Drier, E.;
J. Biol. Chem. 267, 25220-25227, 1992
A/Title: Osteogenic protein-2, A new member of the transforming growth factor-beta superf
A/Reference number: A45056; MUID:93094231; PMID:1160021
A/Accession: A45056
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: nucleic acid
A/Residues: 1-402 <OKZ>
A/Cross-references: UNIPROT:P34820; GB:M97016; NID:G189389; PIDN:AA01360.1; PID:G189390
A/Experimental source: hippocampus

A>Note: sequence extracted from NCBI backbone (NCBI:120189)
C/Genetics:
A:Gene: GDB:BMP8; OP-2
A:Cross-references: GDB:136392
C:Superfamily: inhibin

Alignment Scores:
Pred. No.: 87.7 Length: 402
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x A45056 (1-402)

OY 11 CCGGTCGACCCCGCCCGCT 34
DB 55 ProglyArgProArgProArgAla 62

RESULT 44
E86858
cell division protein FtsZ [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: E86858
R/Bolotin, A.; Winkler, P.; Manger, S.; Dailion, O.; Malarne, K.; Weissensbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp.
A:Reference number: A6625; MUID:21235186; PMID:11337471
A:Accession: E86858
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-417 <STO>
A:Cross-references: UNIPROT:O9C8H2; GB:AE005176; PID:g12724901; PIDN:AAK05967.1; GSPDB:G
A:Experimental source: strain IL1403
C/Genetics:
A:Gene: ftsZ
C:Superfamily: cell division protein ftsZ

Alignment Scores:
Pred. No.: 87 Length: 417
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x E86858 (1-417)

OY 838 CTGACTGAGCGCTCGAGAGCT 861
DB 180 LeuthrgAlaLeuArgGluAla 187

RESULT 45
AG3257
CBS domain containing protein [imported] - Brucella melitensis (strain 16M)
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AG3257
R/DelVecchio, V.G.; Kapratl, V.; Redkar, R.D.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Leese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3257
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-435 <KJF>
A:Cross-references: UNIPROT:Q8FYN8; UNIPROT:Q8FY60; GB:AE008917; PIDN:AAL51226.1; PID:g1
A:Experimental source: strain 16M
C/Genetics:
A:Gene: BME10044

A/Map position: 1

Alignment Scores:
Pred. No.: 86.1 Length: 435
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x AG3257 (1-435)

OY 529 CTGGGAGCCCTCTCCGCCCGCTC 506
DB 138 LeuGlyProLeuSerAlaAlaVal 145

RESULT 46
T07868
DNA binding protein E4/E8BP-1 - tomato
C/Species: Lycopersicon esculentum (tomato)
C/Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C/Accession: T07868
R/Coupe, S.A.; Deikman, J.
Plant J. 11, 1207-1218, 1997
A>Title: Characterization of a DNA-binding protein that interacts with 5' flanking region
A:Reference number: Z16186; MUID:97368850; PMID:9225464
A:Accession: T07868
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-460 <CDU>
A:Cross-references: UNIPROT:O24423; EMBL:U67168; NID:g2342859; PIDN:AA67671.1; PID:g234
A:Experimental source: strain VENT Cherry; fruit
C/Function:
A:Description: may play a role in regulation of the genes E4 and E8, which are coordinat
A>Note: DNA-binding activity increases during fruit ripening

Alignment Scores:
Pred. No.: 85.1 Length: 460
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x T07868 (1-460)

OY 437 AGTGGAGACAGGCTCCCGCCCA 460
DB 140 SerGlyArgGlnAlaProGln 147

RESULT 47
S77558
hypochemical protein s111967 - Synechocystis sp. (strain PCC 6803)
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S77558
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yaeuda,
DNA Res. 3, 109-116, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77558
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-463 <KAN>
A:Cross-references: UNIPROT:P73374; EMBL:D90905; GB:AE001339; NID:g1652360; PIDN:BA11740;
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: hypochemical protein H10333

Alignment Scores:
Pred. No.: 85 Length: 463

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.67% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x S77558 (1-463)

QY 318 CCGCTTGGCGCCAGTGGCGCAG 295
Db 333 ProteulalarginalValArgGln 340

RESULT 48

D64485 TRK system potassium uptake protein homolog - Methanococcus jannaschii

C/Species: Methanococcus jannaschii

C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004

C/Accession: D64485

R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R. Reich, C.I.; Overbeek, R.; Kirschner, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A/Reference number: A64300; MUID:96337999; PMID:8688087

A/Accession: D64485

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-474 <BLU>

A/Cross-references: UNIPROT:Q58880; GB:U67589; GB:L77117; NID:g2826417; PIDN:AA99495.1

C/Genetics:

A/Map position: FOR1455229-1456653

C/Superfamily: potassium uptake protein trkg

Alignment Scores:

Pred. No.: 84.5 Length: 474
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x D64485 (1-474)

QY 44 ATGAGTGCCTGACTACTACGGA 67
Db 111 MetSerlatrptnrtrThnGly 118

RESULT 49

A38340 66k glycoprotein precursor - rabbit

C/Species: Oryctolagus cuniculus (domestic rabbit)

C/Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 09-Jul-2004

C/Accession: A38340

R/Sato, R.; Komine, Y.; Imanaka, T.; Takano, T.

J. Biol. Chem. 265, 2123-2126, 1990

A/Title: Monoclonal antibody EMK1a/212D recognizing site of deposition of extracellular

A/Reference number: A38340; MUID:91065939; PMID:1701177

A/Accession: A38340

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-475 <SAT>

A/Cross-references: UNIPROT:P22458; GB:M55442; GB:J05668; NID:g165037; PIDN:AAA31258.1

C/Superfamily: vitronectin; hemopexin repeat homology; somatomedin B homology

C/Keywords: glycoprotein

F:20-62/Domain: somatomedin B homology <SBH>

F:288-469/Domain: hemopexin repeat homology <PX2>

Alignment Scores:

Pred. No.: 84.5 Length: 475
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.64% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x A38340 (1-475)

QY 484 GGCCGCCAGTGTGTGCGCAGA 507
Db 323 GlyProSerGlyGlyAlaArg 330

RESULT 50

ORHUB1

beta-1-adrenergic receptor - human

C/Species: Homo sapiens (man)

C/Date: 22-Jan-1993 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004

C/Accession: A39911

R/Fieller, T.; Collins, S.; Daniel, K.W.; Caron, M.G.; Lefkowitz, R.J.; Kobilka, B.K.

Proc. Natl. Acad. Sci. U.S.A. 84, 7920-7924, 1987

A/Title: Cloning of the cDNA for the human beta-1-adrenergic receptor.

A/Reference number: A39911; MUID:88068509; PMID:2825170

A/Accession: A39911

A/Molecule type: mRNA

A/Residues: 1-477 <FRI>

A/Cross-references: UNIPROT:P08588; GB:J03019; NID:g178199; PIDN:AA51667.1; PID:g178200

C/Genetics:

A/Gene: GDB:ADRB1; ADRB1R

A/Cross-references: GDB:119654; OMIM:109630

A/Map position: 10q25-10q25

C/Superfamily: vertebrate rhodopsin

C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:58-84/Domain: transmembrane #status predicted <TM1>

F:94-121/Domain: transmembrane #status predicted <TM2>

F:133-154/Domain: transmembrane #status predicted <TM3>

F:177-199/Domain: transmembrane #status predicted <TM4>

F:224-245/Domain: transmembrane #status predicted <TM5>

F:325-346/Domain: transmembrane #status predicted <TM6>

F:358-377/Domain: transmembrane #status predicted <TM7>

F:15/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:

Pred. No.: 84.4 Length: 477
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.64% Indels: 0
DB: 1 Gaps: 0

US-10-030-271-1 (1-909) x ORHUB1 (1-477)

QY 279 GGCCGCCAGACTGTGCGCGCA 302
Db 288 GlyProSerGlyGlyAlaAla 295

Search completed: March 22, 2005, 03:03:56
Job time : 47.3945 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 02:34:17 ; Search time 158.229 Seconds
(without alignments)
5883.642 Million cell updates/sec

Title: US-10-030-271-1
Perfect score: 303
Sequence: 1 atggcgctaccgagctgcac.....tcagtgctgagtgagctgcac 909

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Xgapop 60.0 , Xgapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3224408

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-Q/cg2_1/USFPO_pool_p/US10030271/runat_21032005.153546.26822/app_query.fasta.1.3150
-DB=unafprot_03 -QMT=fastran -SUFFIX=oligo.rup -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030271 @CGN 1.1 578 @runat 21032005.153546.26822 -NCPU=6 -ICPU=3
-NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_crembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	224	73.9	326	1	DEDD2_HUMAN
2	117	38.6	330	1	DEDD2_MOUSE
3	23	7.6	168	2	Q8BRM9
4	14	4.6	369	2	Q919M3
5	13	4.3	218	2	Q919M3
6	13	4.3	318	1	DEDD_HUMAN
7	13	4.3	318	1	DEDD_MOUSE
8	13	4.3	318	1	DEDD_RAT
9	13	4.3	404	2	Q6DHN2
10	11	3.6	243	2	Q6GNZ8
11	10	3.3	218	2	Q91W13
12	10	3.3	235	2	Q98222
13	10	3.3	244	2	Q6DHW2
14	10	3.3	729	2	Q91TW1
15	9	3.0	1461	1	TE18_PRTV
16	9	3.0	110	2	Q8W615
			115	2	Q8C518

17	9	3.0	127	2	Q9Y9Y9
18	9	3.0	148	2	Q67VU8
19	9	3.0	201	2	Q6ZLD4
20	9	3.0	222	2	Q67IV3
21	9	3.0	227	2	Q70G63
22	9	3.0	241	2	Q6MBX0
23	9	3.0	274	2	Q73V06
24	9	3.0	284	2	Q638Y8
25	9	3.0	287	2	Q855F3
26	9	3.0	307	2	Q9RYL0
27	9	3.0	347	2	Q8R353
28	9	3.0	370	1	WNT1_HUMAN
29	9	3.0	373	2	Q72S01
30	9	3.0	374	2	Q6P0P7
31	9	3.0	375	2	Q6PJV4
32	9	3.0	382	2	Q82B67
33	9	3.0	385	2	Q6M204
34	9	3.0	393	2	Q7XJW1
35	9	3.0	395	2	Q8NLT5
36	9	3.0	410	2	Q7MMW4
37	9	3.0	452	2	Q7VUX6
38	9	3.0	454	2	Q8YXA3
39	9	3.0	458	2	Q8BV76
40	9	3.0	465	1	FUMC_SYNPX
41	9	3.0	468	2	Q910P9
42	9	3.0	478	2	Q82HT1
43	9	3.0	516	2	Q6DDP9
44	9	3.0	525	2	Q6K663
45	9	3.0	571	2	Q75EB6
46	9	3.0	629	2	Q871T5
47	9	3.0	636	2	Q6XPZ5
48	9	3.0	709	2	Q8BZN7
49	9	3.0	955	1	T150_HUMAN
50	9	3.0	1099	2	Q7XW40
51	9	3.0	1176	2	Q8PE66
52	9	3.0	1725	2	Q8UW61
53	9	3.0	5068	2	Q63LK7
54	9	3.0	5835	2	Q63LK8
55	8	2.6	27	2	Q6TOB9
56	8	2.6	61	2	Q6W5U5
57	8	2.7	61	2	Q6K2Z8
58	8	2.6	63	2	Q931T4
59	8	2.6	72	2	Q8H7Z8
60	8	2.6	75	2	Q8UHG7
61	8	2.6	79	2	Q8LH71
62	8	2.7	81	2	Q6F611
63	8	2.6	82	2	Q6S600
64	8	2.7	83	2	Q96W43
65	8	2.7	86	2	Q6W4V6
66	8	2.6	91	2	Q9N243
67	8	2.6	91	2	Q9N244
68	8	2.6	91	2	Q9N245
69	8	2.6	91	2	Q9N246
70	8	2.7	91	2	Q6Z1N8
71	8	2.6	93	2	Q9F1P2
72	8	2.6	98	2	Q89XY5
73	8	2.7	100	2	Q9U2F4
74	8	2.6	101	2	Q6UPP9
75	8	2.6	105	2	Q92KNS
76	8	2.6	106	2	Q86V66
77	8	2.7	107	2	Q9LEI7
78	8	2.7	108	2	Q8S5E3
79	8	2.7	108	2	Q7XG46
80	8	2.7	110	2	Q8SDP3
81	8	2.6	113	2	Q8G7A6
82	8	2.6	116	2	Q83W05
83	8	2.7	118	2	Q9CYR2
84	8	2.6	119	2	Q86EP6
85	8	2.6	125	2	Q9BRA0
86	8	2.6	125	2	Q9D2U5
87	8	2.7	128	2	Q69K40
88	8	2.6	129	2	Q653U5
89	8	2.6	134	2	Q7QVZ3

Q9Y9Y9	aeropyrum p
Q67VU8	oryza sativ
Q6ZLD4	oryza sativ
Q67IV3	oryza sativ
Q70G63	anopheles g
Q6MBX0	parachlamy
Q73V06	mycobacteri
Q638Y8	burkholderi
Q855F3	mycobacteri
Q9RYL0	deinococcus
Q8R353	mus musc
WNT1_HUMAN	homo sapien
Q72S01	homo sapien
Q6P0P7	homo sapien
Q6PJV4	homo sapien
Q82B67	streptomyce
Q6M204	cornebacte
Q7XJW1	oryza sativ
Q8NLT5	cornebacte
Q7MMW4	borderella
Q7VUX6	borderella
Q8YXA3	anabaena sp
Q8BV76	mus musc
FUMC_SYNPX	synechococ
Q910P9	streptomyce
Q82HT1	streptomyce
Q6DDP9	xenopus lae
Q6K663	oryza sativ
Q75EB6	ashbya gos
Q871T5	neurospora
Q6XPZ5	triticum ae
Q8BZN7	mus musc
Q92W40	homo sapien
Q7XW40	oryza sativ
Q8PE66	xanthomonas
Q8UW61	oryza sativ
Q63LK7	burkholderi
Q63LK8	burkholderi
Q6TOB9	haptophyte c
Q6W5U5	haptophyte c
Q6K2Z8	oryza sativ
Q931T4	streptomyce
Q8H7Z8	oryza sativ
Q8UHG7	agrobacteri
Q8LH71	oryza sativ
Q6F611	streptococ
Q6S600	oryza sativ
Q96W43	ophiostoma
Q6W4V6	patitrella
Q9N243	presbytis c
Q9N244	nasalis lar
Q9N245	macaca mla
Q9N246	papio hamad
Q6Z1N8	oryza sativ
Q9F1P2	enterococcu
Q89XY5	bradyrhizob
Q9U2F4	caenorhabdi
Q6UPP9	human immu
Q92KNS	rhizobium m
Q86V66	homo sapien
Q9LEI7	hordeum vul
Q8S5E3	oryza sativ
Q7XG46	oryza sativ
Q8SDP3	pseudomonas
Q8G7A6	blifidobacte
Q83W05	thiobacpa r
Q9CYR2	mus musc
Q86EP6	echistosoma
Q9BRA0	homo sapien
Q9D2U5	mus musc
Q69K40	oryza sativ
Q653U5	oryza sativ
Q7QVZ3	giardia lam

C 90 8 2.7 134 2 054030 054030 propionigen
 C 91 8 2.7 134 2 0630Y1 0630Y1 burkholderi
 C 92 8 2.6 138 2 062A50 062A50 burkholderi
 C 93 8 2.6 138 2 0630J1 0630J1 burkholderi
 C 94 8 2.6 142 2 06ERC3 06ERC3 oryza sativ
 C 95 8 2.7 143 2 06MYB1 06MYB1 aspergillus
 C 96 8 2.7 146 2 061M0 061M0 dirosophila
 C 97 8 2.7 146 2 074YH1 074YH1 yersinia pe
 C 98 8 2.7 147 2 092LV3 092LV3 rhizobium m
 C 99 8 2.6 150 2 0981A0 0981A0 rhizobium l
 100 8 2.6 152 2 09DUD0 09DUD0 ct virus. o

 RESULT 1 ALIGNMENTS
 DED2_HUMAN STANDARD; PRT; 326 AA.
 AC 08WKF8; 08NBR2; 08NES1; 08TA48; 096D35;
 DT 05-JUL-2004 (Rel. 44, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE DNA-binding death effector domain-containing protein 2 (DED-containing
 DE protein FLAME-3).
 GN Name=DED2; Synonyms=FLAME3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OC NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND INTERACTION
 RP WITH CASP8.
 RX MEDLINE=21850646; PubMed=11741985; DOI=10.1074/jbc.M10749200;
 RA Roth W., Steiner-Llewen F., Pawlowski K., Godzik A., Reed J.C.;
 RT "Identification and characterization of DED2, a death effector
 RT domain-containing protein.";
 RL J. Biol. Chem. 277:7501-7508(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, SUBCELLULAR
 RP LOCATION, AND INTERACTIONS WITH CASP8 AND GTF3C3.
 RX MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
 RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T.,
 RA Alnemri E.S.;
 RT "Death effector domain-containing proteins DEDD and FLAME-3 form
 RT nuclear complexes with the TP53p102 subunit of human transcription
 RT factor p53.";
 RL Cell Death Differ. 9:439-447(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=22202088; PubMed=12235123; DOI=10.1083/jcb.200112124;
 RA Lee J.C., Schickling O., Stegh A.H., Oshima R.G., Dinsdale D.,
 RA Cohen G.M., Peter M.E.;
 RT "DED2 regulates degradation of intermediate filaments during
 RT apoptosis.";
 RL J. Cell Biol. 158:1051-1066(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Testis; Carcinoma;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Nishimura K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori S., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
 RA Iwida S., Oono Y., Takiguchi S., Watanabe S., Yosida M., Houchi T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Imose N., Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotaka S.,

RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Teraishi Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakeda H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamaguchi M., Watanabe K., Kumagai A., Itakura S., Fukumizu Y.,
 RA Fujimori Y., Komiyama M., Tashiro K., Tanigami A., Fujimura T.,
 RA Oono T., Yamada K., Fujii Y., Ozaki K., Hirose M., Ohnori Y., Okamoto S.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtsuki R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,
 RA Matsumura K., Nakajima T., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Brain, and Lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik A., Farmer A.A., Rubin G.W., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Cassavani T.L., Scheetz T.E.,
 RA Brownstein M.J., Umeda T.B., Toshiyuki S., Carninci P., Prence C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulvaney S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hailton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smolins D.E.,
 RA Schermer A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP INTERACTIONS WITH CASP8 AND CASP10.
 RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
 RA Alcivar A., Hu S., Tang J., Yang X.;
 RT "DED2 and DEDD2 associate with caspase-8/10 and signal cell death.";
 RL Oncogene 22:291-297(2003).
 CC -1- FUNCTION: May play a critical role in death receptor-induced
 CC apoptosis and may target Casp8 and Casp10 to the nucleus. May
 CC regulate degradation of intermediate filaments during apoptosis.
 CC May play a role in the general transcription machinery in the
 CC nucleus and might be an important regulator of the activity of
 CC GTF3C3.
 CC -1- SUBUNIT: Interacts with Casp8, Casp10 and GTF3C3. Homodimerizes
 CC and heterodimerizes with DEDD.
 CC -1- SUBCELLULAR LOCATION: Nuclear, accumulated in subnuclear
 CC structures resembling nucleoli.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=08WKF8-1; Sequence=displayed;
 CC Name=2;
 CC IsoId=08WKF8-2; Sequence=VSP_010312;
 CC -1- TISSUE SPECIFICITY: Expressed in most tissues. High levels were
 CC found in liver, kidney, heart, ovary, spleen, testes, skeletal
 CC muscle and peripheral blood leukocytes. Expression was absent or
 CC low in colon and small intestine. Expression is relatively high in
 CC the tumor cell lines chronic myelogenous leukemia K-562 and the
 CC colorectal adenocarcinoma SW480. Expression is moderate in the lung
 CC cervical carcinoma HeLa, the Burkitt's lymphoma Raji, the lung
 CC carcinoma A549, and the melanoma G361. In contrast, two leukemia
 CC cell lines, HL-60 (promyelocytic leukemia) and MOLT-4
 CC (lymphoblastic leukemia), show relatively low levels.

CC -1- DOMAIN: Interactions with CASP8 and CASP10 are mediated by the DED domain.

CC -1- SIMILARITY: Contains 1 death effector (DED) domain.

CC -1- CAUTION: Ref.5 (AAH13372) sequence differs from that shown due to a frameshift in position 186.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

CC -----

DR EMBL: AF443591; AAL48220.1; -

DR EMBL: AF457575; AAM10835.1; -

DR EMBL: AY125488; AAM95240.1; -

DR EMBL: AK075328; BAC11551.1; -

DR EMBL: BC013372; AAH13372.2; ALT_FRAME.

DR EMBL: BC027930; AAH27930.1; -

DR Genbank: HGNC:24450; DEDD2.

DR H-InvDB: HIX0015171; -

DR InterPro: IPR011029; DEATH_Like.

DR InterPro: IPR01875; DED_Like.

DR Pfam: PF01335; DED; 1.

DR SMART: SM00031; DED; 1.

DR PROSITE: PS50168; DED; 1.

KM Alternative splicing: Apoptosis; DNA-binding; Nuclear protein;

KM Transcription regulation.

FT DOMAIN 25 104 DED.

FT DOMAIN 104 109 Nuclear localization signal (Potential).

FT DOMAIN 155 173 Bipartite nuclear localization signal (Potential).

FT VARSPIC 145 149 Missing (in isoform 2).

FT CONFLICT 27 27 /FTID=VSP_010312.

FT CONFLICT 56 56 H -> N (in Ref. 2).

FT CONFLICT 79 79 A -> G (in Ref. 2).

FT CONFLICT 207 207 C -> G (in Ref. 4).

FT CONFLICT 230 230 C -> R (in Ref. 2).

FT CONFLICT 326 326 Missing (in Ref. 5; AAH13372).

SO SEQUENCE 326 AA; 36178 MM; 3FTBDB307CC870CD CRC64;

Alignment Scores:

Pred. No.:	1,46e-202	Length:	326
Score:	224.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.93%	Indels:	0
	1	Gaps:	0

US-10-030-271-1 (1-909) x DED2_HUMAN (1-326)

QY 238 GAGAGCAACTGGCGCTCTCTGGGCAATCTCTGGCGCTGCTGGCCGCGACGACTGCTG 297

DB 80 Glnuerranleuaglyleuaglyleuaglyleuaglyleuaglyleuaglyleuaglyleu 99

QY 298 CCGGACCTGGCGCGGCG 357

DB 100 Prohlsneulaarglylsargargargargargargargargargargargargargarg 119

QY 358 TCACGCTCTTCAAG 417

DB 120 Sererserserserserserserserserserserserserserserserserserserser 139

QY 418 GCAAAATTCAGCGAGGTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477

DB 140 Alaanserserglnglnglnglnglnglnglnglnglnglnglnglnglnglnglnglng 159

QY 478 AGTGGGGGCGCGCGAGGT 537

DB 160 Serarglgargprosergylgylalargargargargargargargargargargargarg 179

QY 538 CAGCAGCAGTGAAGCCCGCGCAGACCTTCTCTGAAGGCAAGTGAAGTGAAGTGAAGTGA 597

DB 180 GlnlnglnsergluproalaargprosergylgylvalthrCysasp1learg 199

QY 598 CTCGGGGTTCAGCAGAGTACTGCGAGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657

DB 200 LeuargylalarglaaglylrgCysgluhlselProalaleuaglnglnglylaalaser 219

QY 658 CGCGCGCCCGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717

DB 220 ArgargProglalaleuAlaarglnglnleuaspvalPheglglnAlathAlaValleu 239

QY 718 CGCTCAAGGACCTGGGCTGT 777

DB 240 ArgserargAspleuglyserValaCysasp1lelyspeserGluleserlyleu 259

QY 778 GAGCGCTTCGCGCGCGACTACTGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837

DB 260 AspAlaPhearglylsarglylrgleuSerglylaaleuueuglnAlaleuarglylaPhe 279

QY 838 CTGACTGAGCGCTTCGCGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 897

DB 280 LeuthrglnAlaleuargGlnAlaValGlyargGlnAlaValaArgleuValserVal 299

QY 898 GATGAGGCTGAC 909

DB 300 AspGlnAlaAep 303

RESULT 2

DED2_MOUSE

ID DED2_MOUSE STANDARD; PRT; 330 AA.

AC Q8QZV0; Q8JZV1;

DT 05-JUL-2004 (Rel. 44, Created)

DT 05-JUL-2004 (Rel. 44, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE DNA-binding death effector domain-containing protein 2 (DED-containing protein FLAME-3).

GN Name=Ded2; Synonyms=Flame3;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC STRAIN=C57BL/6J;

RX MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj.cdd/4401038;

RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T., Alnemri E.S.;

RT "Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the TP53p102 subunit of human transcription factor IIIC".

RL Cell Death Differ. 9:439-447(2002).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;

RA Alcivar A., Hu S., Tang J., Yang X.;

RT "DEDD and DEDD2 associate with caspase-8/10 and signal cell death".

RL Oncogene 22:291-297(2003).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Retina;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Dichtenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L., Stopleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Rokhs S.S., Lotuallano N.A., Peters G.J., Abramson R.D., Muliyil S.J., Bock S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W., Viallana D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: May play a critical role in death receptor-induced
CC apoptosis and may target Casp8 and Casp10 to the nucleus. May
CC regulate degradation of intermediate filaments during apoptosis.
CC May play a role in the general transcription machinery in the
CC nucleus and might be an important regulator of the activity of
CC GTF3C3.
CC -1- SUBUNIT: Interacts with Casp8, Casp10 and GTF3C3. Homodimerizes
CC and heterodimerizes with DED (by similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear, accumulated in subnuclear
CC structures resembling nucleoli (by similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8QZV0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8QZV0-2; Sequence=VSP_010313, VSP_010314;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Expression is high in liver, heart, kidney,
CC and testis but low in brain, spleen, lung, and skeletal muscle.
CC -1- DOMAIN: Interactions with Casp8 and Casp10 are mediated by the DED
CC domain (by similarity).
CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: AF545757; AAM10836.1; -;
DR EMBL: AF543541; AAM33179.1; -;
DR EMBL: BC037043; AAM37043.1; -;
DR MGD: MGI:1914629; 2410050E1RfK.
DR InterPro: IPR011029; DEATH_Like.
DR Pfam: PF01335; DED.
DR SMART: SM00031; DED; 1.
DR PROSITE: PS50168; DED; 1.
KW Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
KW Transcription regulation.
FT DOMAIN 25 104 DED.
FT DOMAIN 104 109 Bipartite nuclear localization signal (Potential).
FT DOMAIN 156 174 (Potential).
FT VARSPIC 1 171 Missing (in isoform 2).
FT VARSPIC 172 200 /FtId=VSP_010313.
FT VARSPIC 172 200 RRRAGLAQQHQHQHQEIQRSSSEKATC -> MGQAQQRG
FT FT ERTCPRTARAWQSDMSLS (in isoform 2).
FT FT /FtId=VSP_010314.
SQ SEQUENCE 330 AA; 36786 MW; 889BC4F9E01304B0 CRC64;

Db 191 ArpProseSerIuGlyysValThrCysAspIleArgLeuArgValArgAlaGluTyr 210
Qy 619 TGAGAGCATGGGCGAGCCCTTGAGAGAGGGGCTGGCATCCCGGCGCCCGAGCGTGGCG 678
Db 211 CyeGluHieGlyProAlaLeuGluGluGluGlyValAlaSerArgArgProGluAlaLeuAla 230
Qy 679 CGGACGCTGAGCGCTGTTGGGCGAGGCCACCGAGTGTCTGGCTCAAGGAGCGTGGCTCT 738
Db 231 ArgGluLeuAspValIleGlyGluAlaThrAlaValLeuArgSerArgAspLeuGlySer 250
Qy 739 GTGTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGACCGCTTCTGGGCGACTAC 798
Db 251 ValValCysAspIleLeuSerGluLeuSerGluLeuSerGluLeuSerGluLeuSerGlu 270
Qy 799 CTAGTGGCGCCCTGCTGAGAGCGCTGGCGGCGGTCTCTGACTGAGAGCGCTGAGAG 858
Db 271 LeuSerGlyAlaLeuLeuGluGluAlaLeuArgGlyValAlaPheLeuThrGluAlaLeuArgGlu 290
Qy 859 GCTGTGGGCGGAGAGGCTGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
Db 291 AlaValGlyArgGluAlaValAlaValArgLeuLeuValSerValAspIleAlaAsp 307

RESULT 3
OGBRM9
ID OGBRM9 PRELIMINARY; PRT; 168 AA.
AC OGBRM9;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Mus musculus A83004PM19 product: similar to death effector domain-
DE containing and DNA-binding protein 2, full insert sequence.
OS Name=ded2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,

RA Konno H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kasaiwaigi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:1757-1771(2000).
RP [6]
RC SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Cortex;
RA Aachen J., Alizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hizemoto K., Hiroaka T., Hirozane T.,
RA Horii F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK043908; BAC1699.1; -.
DR MGI; MGI:191629; Dcd2.
DR GO; GO:0042981; P:regulation of apoptosis; TAS.
RM DNA-binding.
SQ SEQUENCE 168 AA; 17242 MW; F75D71DC489D2425 CRC64;
US-10-030-271-1 (1-909) x Q8BRM9 (1-168)
Alignment Scores:
Pred. No.: 2,8e-12 Length: 168
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.59% Indels: 0
DB: 2 Gaps: 0
QY 583 ACCGTGACATCCGGCTCCGGCTTCGACAGATGCTCCGACATGGCCGCTTGAG 642
DB 83 ThrcyAspIleArgLeuArgValArgAlaGluTyrGluHISgLyProAlaLeuGln 102
QY 643 CAGGCGCG 651
DB 103 GlnGlyVal 105
RESULT 4
Q919M3 PRELIMINARY; PRT; 369 AA.
AC Q919M3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Dcd1.
GN Name=dcd1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
RL EMBL; AF232226; AAF66963.1; -.
DR ZFIN; ZDB-GENE-000616-2; dcd1.
DR GO; GO:0005515; P:protein binding; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.

DR InterPro; IPR011029; DEATH like.
DR InterPro; IPR001875; DED.
DR Pfam; PF01335; DED; 1.
DR PROSITE; PS50168; DED; 1.
SQ SEQUENCE 369 AA; 42244 MW; 73B09E9E17BEC247 CRC64;
US-10-030-271-1 (1-909) x Q919M3 (1-369)
Alignment Scores:
Pred. No.: 0.000841 Length: 369
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.62% Indels: 0
DB: 2 Gaps: 0
QY 577 AAAGTACCTGTGATCCCGCTCCGGCTTCGACAGATGAC 618
DB 235 LysValThrCysAspIleArgLeuArgValArgAlaGluTyr 248
RESULT 5
DEDD HUMAN
ID DEDD HUMAN STANDARD; PRT; 318 AA.
AC 075618; O60737;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Death effector domain-containing protein (Death effector domain-
DE containing testicular molecule) (DEDDPro1) (FLDED-1) (KE05).
GN Name=DEDD; Synonyms=DEDDPro1, DEFT;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98447599; PubMed=974341; DOI=10.1093/emboj/17.20.5974;
RA Stegk A.H., Schickling O., Ehret A., Scaffidi C., Peterhaensel C.,
RA Hofmann T.G., Grunert I., Kramer P.H., Peter M.E.;
RT "DEDD, a novel death effector domain-containing protein, targeted to
RT the nucleolus.";
RL EMBO J. 17:5974-5986(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISSUE=Testis;
RX MEDLINE=99049260; PubMed=9832420; DOI=10.1210/en.139.12.4839;
RA Leo C.P., Hsu S.Y., McGe E.A., Salanova M., Hsuen A.J.W.;
RT "DEFT, a novel death effector domain-containing molecule predominantly
RT expressed in testicular germ cells.";
RL Endocrinology 139:4839-4848(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX Pan G.;
RT "FLDED-1, a novel molecule with a DED-like domain.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX Thome M., Tschopp J.;
RT "DEDDPro1, a novel DED-containing protein.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Dendritic cell;
RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;
RT "A novel gene from human dendritic cell.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Bone marrow; and Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Mansura K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., MCEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Ramey J., Helton E., Ketterman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywiński M.I., Skalska U., Smallus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[7]
RN FUNCTION, AND INTERACTIONS WITH KRT8, KRT18 AND CASP3.
RP MEDLINE=222202088; PubMed=122353123; DOI=10.1083/jcb.200112124;
RX Lee J.C., Schickling O., Stegh A.H., Oshima R.G., Dinsdale D.,
RA Cohen G.M., Peter M.E.;
RT "DADD regulates degradation of intermediate filaments during
RT apoptosis.";
RT J. Cell Biol. 158:1051-1066(2002).
[8]
RN INTERACTION WITH GTF3C3.
RP MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
RX Zhang Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T.,
RA Alcivar A., Hu S., Tang J., Yang X.;
RT "DADD and DEDD2 associate with caspase-8/10 and signal cell death.";
RT Oncogene 22:291-297(2003).
RL -1- FUNCTION: A scaffold protein that directs CASP3 to certain
CC substrates and facilitates their ordered degradation during
CC apoptosis. May also play a role in mediating CASP3 cleavage of
CC KRT18. Regulates degradation of intermediate filaments during
CC apoptosis. May play a role in the general transcription machinery
CC in the nucleus and might be an important regulator of the activity
CC of GTF3C3. Inhibits DNA transcription in vitro (By similarity).
CC -1- SUBUNIT: Interacts with CASP8, CASP10, KRT8, KRT18, CASP3 and
CC RADD. Homodimerizes and heterodimerizes with DEDD2.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
CC during C95-mediated apoptosis where it is localized in the mono
CC nucleoli (By similarity). Following apoptosis induction, the mono
CC and/or d tubulin/actin form increases and forms filamentous
CC structures that colocalize with KRT8 and KRT18 intermediate
CC filament network in simple epithelial cells.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O75618-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O75618-2; Sequence=VSE_003846;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
CC testis.
CC -1- PFM: Exists predominantly in a mono- or d tubulin/actin form.
CC -1- SIMILARITY: Contains 1 death effector (DED) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; AF083236; AAC33105.1; -.
DR EMBL; AF100341; AAD16414.1; -.
DR EMBL; AF043733; AAC80280.1; -.
DR EMBL; AJ010973; CAA09445.1; -.
DR EMBL; AF064605; AAC17110.3; -.
DR EMBL; BC016724; AAH16724.1; -.
DR EMBL; BC013910; AAH13910.1; -.
DR Genew; HNGC:2755; DEDD.
DR H-invDB; HIX001231; -.
DR MIM; 606841; -.
DR GO; GO:0005737; C:cytoplasm; ISS.
DR GO; GO:0005730; C:nucleolus; ISS.
DR GO; GO:0003677; F:DNA binding; ISS.
DR GO; GO:000617; P:induction of apoptosis; ISS.
DR GO; GO:0008625; P:induction of apoptosis via death domain rec. .; TAS.
DR GO; GO:0016481; P:negative regulation of transcription; ISS.
DR GO; GO:0007283; P:spermatogenesis; TAS.
DR InterPro; IPR011029; DEATH_like.
DR InterPro; IPR01875; DED.
DR Pfam; PF01335; DED; 1.
DR SMART; SM00031; DED; 1.
DR PROSITE; PS50168; DED; 1.
KW Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
  Repressor; Transcription regulation.
FT DOMAIN 25 103 DED.
FT VARSPIC 194 194 DED.
FT D -> GREGFORSHLEGYKELLGHMAVYAIQY (in
  isoform 2).
FT FT /Frid=VSP_003846.
FT SEQUENCE 318 AA; 36794 MW; PF9D5FPB61F6B6 CRC64;
  P -> L (in Ref. 5).

Alignment Scores:
Pred. No.: 0.00757 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
DB: 1 Gaps: 0

US-10-030-271-1 (1-909) x DEDD_HUMAN (1-318)

QY 583 ACTGTGATCCGGCTCCGGGTTCGACGAGTACTGC 621
DB 192 ThrcysaspIleArgLeuArgValArgAlaGluIrrCys 204
-----
DEDD_MOUSE STANDARD; PRT; 318 AA.
ID DEDD_MOUSE
AC Q92IL3; Q7T0H8; Q9R227;
  28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Death effector domain-containing protein (DEDDProl).
GN Name=Dedd;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCB1_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
  SPECIFICITY.
RX MEDLINE=98447599; PubMed=9774341; DOI=10.1093/emboj/17.20.5974;
RA Stegh A.H., Schickling O., Ehret A., Scaffidi C., Peterhaensel C.,
  Hofmann T.G., Grunnt I., Kraemer P.H., Peter M.B.;
RT "DEDD, a novel death effector domain containing protein, targeted to
  the nucleolus.";
RL EMOB J. 17:5974-5986 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Thome M., Tschopp J.;
RT "DEDDProl, a novel DED-containing protein.";
```

RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FWB/N-3; TISSUE=Colon, and Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin J., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Uedin T.A., Toshlyuk S., Carrinci P., Prange C.,
 RA Rana S.S., Loquillano N.B., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villation D.K., Muzny D.M., Sodegren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: A scaffold protein that directs CASP3 to certain
 CC substrates and facilitates their ordered degradation during
 CC apoptosis. May also play a role in mediating CASP3 cleavage of
 CC KRT18. Regulates degradation of intermediate filaments during
 CC apoptosis. May play a role in the general transcription machinery
 CC in the nucleus and might be an important regulator of the activity
 CC of GTF3C3 (by similarity). Inhibits DNA transcription in vitro.
 CC -1- SUBUNIT: Interacts with CASP8, CASP10, KRT8, KRT18, CASP3 and
 CC FADD. Homodimerizes and heterodimerizes with DED2 (by
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
 CC during CD95-mediated apoptosis where it is localized in the
 CC nucleoli. Following apoptosis induction, the mono and/or
 CC dimerization form increases and forms filamentous structures
 CC that colocalize with KRT8 and KRT18 intermediate filament network
 CC in simple epithelial cells (by similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- PFM: Exists predominantly in a mono- or dimerized form.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AJ011386; CAA09604.1; -
 CC EMBL: AF100342; AAD16415.1; -
 CC EMBL: BC023668; AAH23668.1; -
 CC EMBL: BC054445; AAH54445.1; -
 CC MGI: MGI:1333874; Dedd.
 CC GO: GO:0005737; C:cytoplasm; IDA.
 CC GO: GO:0005730; C:nucleolus; IDA.
 CC GO: GO:0003677; F:DNA binding; IDA.
 CC GO: GO:0006917; P:induction of apoptosis; IDA.
 CC GO: GO:0016481; P:negative regulation of transcription; IDA.
 CC InterPro: IPR011029; DEDATH_like.
 CC InterPro: IPR018753; DED.
 CC Pfam: PF01335; DED; 1.
 CC SMART: SM00031; DED; 1.
 CC PROSITE: PS50168; DED; 1.
 CC Apoptosis; DNA-binding; Nuclear protein; Repressor;
 CC Transcription regulation.
 CC DOMAIN 25 103 DED.
 CC CONFLICT 237 237 K -> N (in Ref. 2).
 CC CONFLICT 316 316 A -> V (in Ref. 3; AAH54445).

SO SEQUENCE 318 AA; 36805 MW; C9A31DFC4C0E57CA CRC64;
 Alignment Scores:
 Pred. No.: 0.00757 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 4.29% Indels: 0
 DB: 1 Gaps: 0
 US-10-030-271-1 (1-909) x DEDD_MOUSE (1-318)
 QY 583 ACCTGTGACATCGCGCTCGGCTTCAGACAGACTACGC 621
 DB 192 Thrcysapleleatglaargvalargalacturrcys 204
 RESULT 7
 ID DEDD_RAT STANDARD; PRT; 318 AA.
 AC Q922K0;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Death effector domain-containing protein (death effector domain-
 DE containing testicular molecule).
 GN Name=Dedd; Synonyms=Def1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE-Testis;
 RX MEDLINE=99048260; PubMed=9832420; DOI=10.1210/en.139.12.4839;
 RA Leo C.P., Hsu S.Y., McGe E.A., Salanova M., Hsuen A.J.W.;
 RT "DEFT, a novel death effector domain-containing molecule predominantly
 RT expressed in testicular germ cells."
 RL Endocrinology 139:4839-4848(1998).
 CC -1- FUNCTION: A scaffold protein that directs CASP3 to certain
 CC substrates and facilitates their ordered degradation during
 CC apoptosis. May also play a role in mediating CASP3 cleavage of
 CC KRT18. Regulates degradation of intermediate filaments during
 CC apoptosis. May play a role in the general transcription machinery
 CC in the nucleus and might be an important regulator of the activity
 CC of GTF3C3. Inhibits DNA transcription in vitro (by similarity).
 CC -1- SUBUNIT: Interacts with CASP8, CASP10, KRT8, KRT18, CASP3 and
 CC FADD. Homodimerizes and heterodimerizes with DED2 (by
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
 CC during CD95-mediated apoptosis where it is localized in the
 CC nucleoli. Following apoptosis induction, the mono and/or
 CC dimerization form increases and forms filamentous structures
 CC that colocalize with KRT8 and KRT18 intermediate filament network
 CC in simple epithelial cells (by similarity).
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
 CC testis. Within the testis, highly expressed in germ cells but not
 CC expressed in Sertoli cells.
 CC -1- DEVELOPMENTAL STAGE: First detected in 20-day-old animals. Reaches
 CC a peak at 30 days.
 CC -1- PFM: Exists predominantly in a mono- or dimerized form.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AF053362; AAC00287.1; -
 CC GO: GO:0005737; C:cytoplasm; ISS.
 CC GO: GO:0005730; C:nucleolus; ISS.

DR GO; GO:0003677; F:DNA binding; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.
DR GO; GO:0016481; P:negative regulation of transcription; ISS.
DR InterPro; IPR01029; DEATH_like.
DR InterPro; IPR001875; DED.
DR Pfam; PF01335; DED; 1.
DR SMART; SM00031; DED; 1.
DR PROSITE; PS0168; DED; 1.
KW Apoptosis; DNA-binding; Nuclear protein; Repressor;
KW Transcription regulation.
FT DOMAIN 25 103 DED.
SQ SEQUENCE 318 AA; 36847 MW; B8751791F66A03DE CRC64;
DB: 1
Alignment Scores:
Pred. No.: 0.00757 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
Gaps: 0
DB: 1
US-10-030-271-1 (1-909) x DEDD_RAT (1-318)
QY 583 ACCTGACATCGGCTCGGTTTCAGCAGACTGTC 621
DB 192 ThCyaspilleargleuargValArgAlaGluIuTcyS 204
RESULT 8
Q6DNH2 PRELIMINARY; PRT; 404 AA.
AC Q6DNH2; 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:92202.
GN Name=zgc:92202;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hailton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC075935; AAH75935.1; -;
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.

DR InterPro; IPR01029; DEATH_like.
DR InterPro; IPR001875; DED.
DR Pfam; PF01335; DED; 1.
DR SMART; SM00031; DED; 1.
DR PROSITE; PS0168; DED; 1.
SQ SEQUENCE 404 AA; 44888 MW; 3C949DBB3B07B81A CRC64;
DB: 2
Alignment Scores:
Pred. No.: 0.00735 Length: 404
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
Gaps: 0
DB: 2
US-10-030-271-1 (1-909) x Q6DNH2 (1-404)
QY 583 ACCTGACATCGGCTCGGTTTCAGCAGACTGTC 621
DB 267 ThCyaspilleargleuargValArgAlaGluIuTcyS 279
RESULT 9
Q6GNZ8 PRELIMINARY; PRT; 243 AA.
AC Q6GNZ8; 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC80767 protein.
GN Name=MGC80767;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loguallano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hailton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buttefield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative";
RL Dev. Dyn. 225:384-391(2002).
[3]
RN SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S., Gerhart D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR001875; DED.
DR Pfam: PF01335; DED; 1.
DR PROSITE: PSS0168; DED; 1.
SQ SEQUENCE 244 AA; 28051 MW; 043225AEASD79527 CRC64;

Alignment Scores:
Pred. No.: 5.4 Length: 244
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.30% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q6DHV2 (1-244)

Qy 199 GAGCTCCTGCTGAGCTGAGCCGCGG 228
Db 113 Gtuleuleuleugtuleugtargargly 122

RESULT 13
Q91TW1 PRELIMINARY; PRT; 729 AA.
AC Q91TW1;
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 12.
OS Tupaiid herpesvirus 1 (strain 1) (TUVH-1) (Herpesvirus tupai (strain 1)).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae;
OX NCBI_TaxID=10397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RX MEDLINE=21211637; PubMed=11312357;
RX DOI=10.1128/JVI.75.10.4854-4870.2001;
RA Bahr U., Darai G.;
RT "Analysis and characterization of the complete genome of tupai (tree shrew) herpesvirus.";
RL J. Virol. 75:4854-4870(2001).
DR EMBL: AF281817; AAK57026.1; -.
DR Pfam: PF02393; US22; 1.
SQ SEQUENCE 729 AA; 81191 MW; D5FDE57AAF3EAB5 CRC64;

Alignment Scores:
Pred. No.: 4.73 Length: 729
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q91TW1 (1-729)

Qy 224 CGGCGCTCCAGCTCCAGCAGAGCTCTAGG 195
Db 465 ArgArgSerSerSerSerArgSerSerArg 474

RESULT 14
IE18_PRIVIF
ID IE18_PRIVIF STANDARD; PRT; 1461 AA.
AC P1675;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Immediate-early protein IE180.
GN Name=IE1;
OS Pseudorabies virus (strain Indiana-Punkhauser / Becker)
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=31523;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=89315207; PubMed=2546124;
RA Cheung A.K.;
RT "DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus.";
RL Nucleic Acids Res. 17:4637-4646(1989).
RN [2]
RP REVISIONS.
RA Cheung A.K.;
RL Submitted (NOV-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This IE protein is a multifunctional protein capable of migrating to the nucleus, binding to DNA, trans-activating other viral genes, and autoregulating its own synthesis.
CC -1- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -1- PTM: A long stretch of serine residues may be a major site of phosphorylation.
CC -1- SIMILARITY: Belongs to the herpesviruses ICP4/IE140/IE180 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15120; CA33214.1; -.
DR PIR: S04713; EDBEIRF.
DR InterPro: IPR005205; Herpes ICP4_C.
DR InterPro: IPR005206; Herpes ICP4_N.
DR Pfam: PF03585; Herpes_ICP4_C_1.
DR Pfam: PF03584; Herpes_ICP4_N_1.
KW DNA-binding; Early protein; Nuclear protein; Phosphorylation;
KM Trans-acting factor; Transcription regulation.
FT DOMAIN 390 405 Poly-Ser.
FT DOMAIN 958 966 Poly-Ser.
SQ SEQUENCE 1461 AA; 149833 MW; 7F31E7ABE403B208 CRC64;

Alignment Scores:
Pred. No.: 4.35 Length: 1461
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.33% Indels: 0
DB: 1 Gaps: 0

US-10-030-271-1 (1-909) x IE18_PRIVIF (1-1461)

Qy 536 GGTGCGGCTGGGGCCCTCTCCGCCGCCCT 507
Db 1047 GlyAlaIadIyAlaProLeuAArgArg 1056

RESULT 15
Q8W615
ID Q8W615 PRELIMINARY; PRT; 110 AA.
AC Q8W615;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein orf42.
GN Name=orf42;
OS Sinorhizobium meliloti phage PBC5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.
OX NCBI_TaxID=179237;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulmeister S.A., Krol J.E., Vorhoeelter F.-J., Skornupka A.M., Lotz W.;
RT "Sequence of the genome of sinorhizobium meliloti bacteriophage PBC5.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RA Schulteier S.A., Krol J.E., Vorhoeiter F.-J., Skorupska A.M.;
RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF448724; AAL49638.1; -
KW Hydrothermal protein.
SQ SEQUENCE 110 AA; 12194 MW; 1EF7398F8CF302E8 CRC64;

Alignment Scores:
Pred. No.: 52.6 Length: 110
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q8W615 (1-110)

OY 215 AGCTCCAGCAGAGCTCTAGCGCCCTG 189
Db 100 SerSerSerArgSerSerArgProLeu 108

RESULT 16
O8C518 PRELIMINARY; PRT; 115 AA.
AC O8C518;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430539E12 product:hypothetical protein, full
insert sequence.
GN Name=2610204G22R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RA The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=20530913; PubMed=11075861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Kono H., Akiyama U., Nishi K., Kiteunai T., Tashiro H., Itoh M.,
Saito N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,

RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaishiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system 384-format
sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohnato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AK078249; BAC37192.1; -
DR MGD; MGI:1917698; 2610204G22R1k.
KW Hypothetical protein.
SQ SEQUENCE 115 AA; 12569 MW; BBEB1C246D9D4C52 CRC64;

Alignment Scores:
Pred. No.: 52.3 Length: 115
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x O8C518 (1-115)

OY 114 CGAGTGCAGCTGAGCTCTAGCGCCTT 140
Db 95 ArgValArgAlaGlyAlaProGlyLeu 103

RESULT 17
Q9Y9Y9 PRELIMINARY; PRT; 127 AA.
AC Q9Y9Y9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE2150.
GN OrderedLocustNames=APE2150;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hano Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoya K., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Punahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kashiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000063; BAA81161.1; -
DR PIR; A72522; A72522.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 127 AA; 13517 MW; 1C2BD0E67B34434C CRC64;

Alignment Scores:

RA Whitehead S., Yeats C., Bartell B.G., Oyston P.C.F., Parkhill J.,
 RT "Genomic plasticity of the causative agent of melioidosis,
 RL Burkholderia pseudomallei.";
 DR Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
 SO EMBL: BX571965; CAH36185.1; -.
 SQ SEQUENCE 284 AA; 30161 MW; 491795AFE38AA06 CRC64;

Alignment Scores:
 Pred. No.: 46.9 Length: 284
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q63SY8 (1-284)

QY 694 ACAGCTCAGCTGCGGCGAGCGCT 668
 Db 40 ThraGrProAlaAlaProAlaPro 48

RESULT 25
 Q855F3 PRELIMINARY; PRT; 287 AA.
 AC Q855F3; 287 AA.
 DT 01-JUN-2003 (TReMBLrel. 24, Created)
 DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE G857.
 OS Mycobacteriophage Cheb
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxID=205868;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Madsworth C.,
 RA Lewis J.A., Jacobs-Sera D., Falbo U., Gross U., Pannunzio N.R.,
 RA Brucher W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
 RA Hatfull G.F.;
 RT "Origins of highly mosaic mycobacteriophage genomes.";
 RL Cell 113:171-183(2003)
 DR EMBL: AY129330; ANA12455.1; -
 DR GO: GO:0003700; F:transcription factor activity; IEA.
 DR GO: GO:0006355; P:regulation of transcription; DNA-dependent; IEA.
 DR InterPro: IPR002197; HTH_Fis.
 DR PRINTS: PR01590; HTHFIS.
 SQ SEQUENCE 287 AA; 28977 MW; 9AFA47531C54D08B CRC64;

Alignment Scores:
 Pred. No.: 46.8 Length: 287
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.97% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q855F3 (1-287)

QY 215 TGGAGCGCGGCGGCGAGCTGCGGCGAGA 241
 Db 158 TtpSerAlaAlaGlySerAlaAlaArg 166

RESULT 26
 Q89YLO PRELIMINARY; PRT; 307 AA.
 AC Q89YLO; 307 AA.
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE Hypothetical protein DRA0302.
 GN OrderedLocusNames=DRA0302;
 OS Deinococcus radiodurans.

OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI / ATCC 13939 / DSM 20539 / NCIB 9279;
 RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L.A., Ueberback T.R., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans RI.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001863; AAP12516.1; -.
 DR PIR: C75584; C75584.
 DR TIGR; DRA0302; -.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 307 AA; 32062 MW; EED6754B565DC8 CRC64;

Alignment Scores:
 Pred. No.: 46.4 Length: 307
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q89YLO (1-307)

QY 226 CGCGGCGCTCCAGCTCCAGCGAGCT 200
 Db 278 ArgGlyAlaProAlaProAlaGlyAla 286

RESULT 27
 Q8R353 PRELIMINARY; PRT; 347 AA.
 AC Q8R353; 347 AA.
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Ttrp3 protein.
 GN Name=Thrap3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muljany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyzanski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strauberg R.;
RU Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026580; AAH2580.1; -.
DR MGD; MGI:2442637; Thrap3.
DR GO; GO:0000119; C:mediator complex; IDA.
DR GO; GO:0006366; P:transcription from Pol II promoter; IC.
SQ SEQUENCE 347 AA; 38293 MW; C685A3C2394FADa6 CRC64;

Alignment Scores:
Pred. No.: 45.8 Length: 347
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0

US-10-030-271-1 (1-909) x Q8R353 (1-347)

Qy 224 GCGGCTTCAGCTCCAGAGAGCTCT 198
Db 155 ArgArgSerSerSerArgSerSer 163

RESULT 28
WNT1_HUMAN
ID WNT1_HUMAN STANDARD; PRT; 370 AA.
AC P04628;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Wnt-1 proto-oncogene protein precursor.
GN Name=WNT1; Synonyms=INT1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86055728; PubMed=2998762;
RA van Ooyen A.; Kwee V.; Nusse R.;
RT "The nucleotide sequence of the human int-1 mammary oncogene;
RT evolutionary conservation of coding and non-coding sequences.";
RL EMBO J. 4:2905-2909(1985).
CC -1- FUNCTION: Ligand for members of the frizzled family of seven
CC transmembrane receptors. Probable developmental protein. May be a
CC signaling molecule important in CNS development. Is likely to
CC signal over only few cell diameters.
CC -1- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC extracellular matrix.
CC -1- SIMILARITY: Belongs to the Wnt family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X03072; CAA26874.1; -.
DR PIR; A24674; TVHUT1.
DR Genew; HGNC:12774; WNT1.
DR MIM; 164820; -.
DR GO; GO:0007417; P:central nervous system development; TAS.
DR GO; GO:0006653; P:morphogenesis; TAS.
DR InterPro; IPR005817; Wnt.
DR InterPro; IPR009139; Wnt1.
DR InterPro; IPR005816; Wnt_guthfactor.
DR Pfam; PF00110; wnt; 1.
DR PRINTS; PRO1841; WNT1PROTEIN.

DR PRINTS; PRO1349; WNTPROTEIN.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Proto-oncogene; Signal;
KW Wnt signaling pathway.
FT SIGNAL 1 27 Potential.
FT CHAIN 28 370 Wnt-1 proto-oncogene protein.
FT CARBOHYD 29 29 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 316 316 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 346 346 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 359 359 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 370 AA; 40981 MW; F7E811DA12E173F CRC64;

Alignment Scores:
Pred. No.: 45.4 Length: 370
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0

US-10-030-271-1 (1-909) x WNT1_HUMAN (1-370)

Qy 243 GCTCTGCCGCGACTGCGCGCGCTC 217
Db 18 AlaleuAlaleuProAlaleu 26

RESULT 29
Q725U1
ID Q725U1 PRELIMINARY; PRT; 373 AA.
AC Q725U1;
DT 01-OCT-2003 (TRENDSrel. 25, Created)
DT 01-OCT-2003 (TRENDSrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENDSrel. 25, Last annotation update)
DE THRAP3 protein (Fragment).
GN Name=THRAP3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RC Strauberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.;
RA Ditschenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stadler M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;
RA Brownstein M.J.; Udell T.B.; Tohiyoki S.; Carninci P.; Prange C.;
RA Rana S.S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullaly S.J.;
RA Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.;
RA Richards S.; Morley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Huliy S.W.;
RA Villalón D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.;
RA Fahey J.; Helton E.; Kettman M.; Madan A.; Rodriguez S.; Sanchez A.;
RA Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.;
RA Blakeley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.;
RA Roldrequez A.C.; Grimwood J.; Schmutz J.; Myers R.M.; Butterfield Y.S.;
RA Krzywinski M.I.; Skaleja U.; Smalhus D.E.; Schnerch A.; Schein J.E.;
RA Jones S.J.; Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strauberg R.;
RU Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054046; AAH54046.1; -.
DR NON_TER 373 373
SQ SEQUENCE 373 AA; 41600 MW; 05EA5A7DCFP97F9B CRC64;

Alignment Scores:

Pred. No.: 45.4
 Score: 9.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 3.00%
 DB: 2
 Gaps: 0

US-10-030-271-1 (1-909) x Q725U1 (1-373)

Qy 224 CGGCGCTCCAGCTCCAGAGAGCTCT 198

Db 155 ArgArgSerSerSerSerArgSerSer 163

RESULT 30

Q6POP7 PRELIMINARY; PRT; 374 AA.

AC 06POP7; MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

DT 05-JUL-2004 (TRENBLREL. 27, Created)

DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)

DE THRAP3 protein (Fragment).

GN Name=THRAP3;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RA [1]

RA SEQUENCE FROM N.A.

RC TISSUE=Lymph;

RA STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uedini T.B., Toshitsuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RN SEQUENCE FROM N.A.

RC TISSUE=Lymph;

RA STRAUSBERG R.

RU Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC065519; AAH65519.1; -.

FT NON TER 374

SQ SEQUENCE 374 AA; 41729 MW; AF05BA5A75EFP97F CRC64;

Alignment Scores:

Pred. No.: 45.3
 Score: 9.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 3.00%
 DB: 2
 Gaps: 0

US-10-030-271-1 (1-909) x Q6POP7 (1-374)

Qy 224 CGGCGCTCCAGCTCCAGAGAGCTCT 198

Db 155 ArgArgSerSerSerSerArgSerSer 163

RESULT 31

Q6PUV4 PRELIMINARY; PRT; 375 AA.

AC 06PUV4; MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

DT 05-JUL-2004 (TRENBLREL. 27, Created)

DT 05-JUL-2004 (TRENBLREL. 27, Last sequence update)

DE THRAP3 protein (Fragment).

GN Name=THRAP3;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RA [1]

RA SEQUENCE FROM N.A.

RC TISSUE=Placenta; and Skin;

RA STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uedini T.B., Toshitsuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RN SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA STRAUSBERG R.

RU Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RN SEQUENCE FROM N.A.

RC TISSUE=Skin;

RA STRAUSBERG R.

RU Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC010997; AAH10997.1; -.

FT EMBL; BC002501; AAH02501.1; -.

FT NON TER 375

SQ SEQUENCE 375 AA; 41857 MW; 2B6F05BA5A75EFP9 CRC64;

Alignment Scores:

Pred. No.: 45.3
 Score: 9.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 3.00%
 DB: 2
 Gaps: 0

US-10-030-271-1 (1-909) x Q6PUV4 (1-375)

Qy 224 CGGCGCTCCAGCTCCAGAGAGCTCT 198

Db 155 ArgArgSerSerSerSerArgSerSer 163

RESULT 32

Q82B67 PRELIMINARY; PRT; 382 AA.

AC 082B67; MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

DT 01-JUN-2003 (TRENBLREL. 24, Created)

DT 01-JUN-2003 (TRENBLREL. 24, Last sequence update)

```

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=SAV5838;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomycetes.
OC NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531 (2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=2147403; PubMed=11572948; DOI=10.1073/pnas.21143198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
DR EMBL; AP005044; BAC73550.1; -.
KM Complete proteome.
SQ SEQUENCE 382 AA; 41611 MW; FBC5D6358F1FA4D7 CRC64;

Alignment Scores:
Pred. No.: 45.2 Length: 382
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
Gaps: 0
DB: 2

US-10-030-271-1 (1-909) x Q82B67 (1-382)

QY 660 GCGGCCCGAGCGCTGGCGCGCGAGCT 666
Db 133 AlaAlaProGlyAlaGlyAlaAlaAla 141

RESULT 33
Q6M204 PRELIMINARY; PRT; 385 AA.
AC Q6M204;
ID Q6M204;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE BETA-N-ACETYLGLUCOSAMINIDASE (EC 3.2.1.52).
GN Name=nagA2; OrderedLocustNames=cg3158;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteriia; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OC NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goessmann A., Hartmann M., Huttmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahm H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25 (2003).
DR EMBL; BX927156; CAF20877.1; -.

```

```

DR GO; GO:0004563; F:beta-N-acetyl]hexosaminidase activity; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001764; Glyco_hydro_3N.
DR Pfam; PF00933; Glyco_hydro_3; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 385 AA; 40473 MW; 367F80C810D53849 CRC64;

Alignment Scores:
Pred. No.: 45.2 Length: 385
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
Gaps: 0
DB: 2

US-10-030-271-1 (1-909) x Q6M204 (1-385)

QY 844 GAGGCCCTGCGAGAGCGTGGCGCG 870
Db 107 GluAlaLeuArgGluAlaValGlyArg 115

RESULT 34
Q7XJW1 PRELIMINARY; PRT; 393 AA.
ID Q7XJW1;
AC Q7XJW1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE OSJNB40016002.9 protein.
GN Name=OSJNB40016002.9;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OC NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=12447439; DOI=10.1038/nature01183;
RX Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Fan Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Pan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320 (2002).
DR EMBL; AL606588; CA805999.3; -.
DR HSSP; P00909; IUCM.
DR Grimeye; Q7XJW1; -.
DR GO; GO:0004425; F:indole-3-glycerol-phosphate synthase activity; IEA.
DR GO; GO:0006568; P:tryptophan metabolism; IEA.
DR InterPro; IPR003009; FMN enzyme.
DR InterPro; IPR001468; IGPS.
DR InterPro; IPR011060; RibP_bind_barrel.
DR Pfam; PF00218; IGPS; 1.
DR ProDom; PD001511; IGPS; 1.
DR PROSITE; PS00614; IGPS; 1.
SQ SEQUENCE 393 AA; 42409 MW; 80B9F1860BDE8547 CRC64;

Alignment Scores:
Pred. No.: 45.1 Length: 393
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
Gaps: 0
DB: 2

US-10-030-271-1 (1-909) x Q7XJW1 (1-393)

```

QY 301 GCGGACGAGCTGCTGGCGGCCAGCA 275
 |||||
 Db 6 AlalalalaglyargglyglyProAla 14

RESULT 35
 Q8NLT5 PRELIMINARY; PRT; 395 AA.
 AC Q8NLT5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DE Beta-glucosidase-related glycosidases (EC 3.2.1.30).
 GN OrderedLocustNames=Cgl2852;
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 DB NCBI_Taxid=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
 RA Nakagawa S.;

RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.",
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF005283; BAC00246.1; -
 DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
 DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
 DR InterPro; IPR01764; Glyco_hydro_3N.
 DR Pfam; PF00933; Glyco_hydro_3; 1.
 KM Complete proteome; Glycosidase; Hydrolase.
 SQ SEQUENCE 395 AA; 41554 MW; FBF35F878C4F06AD CRC64;

Alignment Scores:
 Pred. No.: 45 Length: 395
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.97% Indels: 0
 DB: NCBI_Taxid=520; Gaps: 0

US-10-030-271-1 (1-909) x Q8NLT5 (1-395)

QY 844 GAGGCGCTGCGAGGCGCTGTGGCGCG 870
 |||||
 Db 117 Glualalalauargglyalalalalarg 125

RESULT 36
 Q7MMW4 PRELIMINARY; PRT; 410 AA.
 AC Q7MMW4;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DE Putative integral membrane protein.
 GN OrderedLocustNames=BB1276;
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 DB NCBI_Taxid=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RA MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Houlden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jørgels K.,
 RA Leathes S., Moule S., Norberczak H., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640440; CAE31774.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; F:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR PROSITE; PS50850; MFS; 1.
 KM Complete proteome.
 SQ SEQUENCE 410 AA; 40657 MW; 135C99999DC40AF0 CRC64;

Alignment Scores:
 Pred. No.: 44.8 Length: 410
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.97% Indels: 0
 DB: NCBI_Taxid=520; Gaps: 0

US-10-030-271-1 (1-909) x Q7MMW4 (1-410)

QY 657 CCGGCGGCGCCGCTGGCGGCCGCA 683
 |||||
 Db 10 ProAlalalProgllyalalalAla 18

RESULT 37
 Q7VUX6 PRELIMINARY; PRT; 452 AA.
 AC Q7VUX6;
 DT 01-OCT-2003 (TREMBlrel. 25, Created)
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
 DE Putative integral membrane protein.
 GN OrderedLocustNames=BP2940;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 DB NCBI_Taxid=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCIC 13251;
 RA MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Houlden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jørgels K.,
 RA Leathes S., Moule S., Norberczak H., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders S., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;

RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.",
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640419; CAE43212.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; F:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR PROSITE; PS50850; MFS; 1.
 KM Complete proteome.
 SQ SEQUENCE 452 AA; 45805 MW; A369C2AC3B174191 CRC64;

Alignment Scores:
 Pred. No.: 44.3 Length: 452
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.97% Indels: 0
 DB: NCBI_Taxid=520; Gaps: 0

```

US-10-030-271-1 (1-909) x Q7VUX6 (1-452)
QY 657 CCGCGCCGCCAGCGCTGCGCGGCA 683
DB 54 ProAlaAlaProG|YAlaG|YAlaAla 62
RESULT 38
Q8YXA3
ID Q8YXA3 PRELIMINARY; PRT; 454 AA.
AC Q8YXA3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE AL1312 protein.
GN OrderedLocusNames=alr1312;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxId=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriuchihara M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003585; BAB73269.1; -.
DR PIR; AE1970; AE1970.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:000506; F: iron ion binding; IEA.
DR Pfam; PF04055; Radical_SAM; 1.
DR Pfam; PF01938; TRAM; 1.
DR Pfam; PF00919; UPF0004; 1.
DR SMART; SM00729; ELP3; 1.
DR TIGRFAMs; TIGR01574; miaB-methylase; 1.
DR TIGRFAMs; TIGR00089; UPF0004; 1.
DR PROSITE; PS01276; TRAM; 1.
DR PROSITE; PS01276; UPF0004; 1.
DR Complete proteome.
SQ SEQUENCE 454 AA; 51509 MW; 8606FD0B3D7BDBC CRC64;

Alignment Scores:
Pred. No.: 44.3 Length: 454
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
Gaps: 0
DB: 2

US-10-030-271-1 (1-909) x Q8YXA3 (1-454)
QY 528 AGCCGACCCGACGAGCTGACGACC 554
DB 175 SerArgThrProAlaAlaValArgAla 183
RESULT 39
Q8BV76
ID Q8BV76 PRELIMINARY; PRT; 458 AA.
AC Q8BV76;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched
DE library, clone:A33062C13 product:similar to THYROID HORMONE RECEPTOR-
DE ASSOCIATED PROTEIN COMPLEX COMPONENT TRAP150 (Fragment).
GN Name=Thrap3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055550;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuura T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachiaki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
RA Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Karukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura T.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK079601; BAC37696.1; -.
DR MGD; MGI:2442637; Thrap3.
DR GO; GO:0000119; C: mediator complex; IEA.
DR GO; GO:0006366; P: transcription from Pol II promoter; IC.
KW Receptor.
FT NON TER 458
SQ SEQUENCE 458 AA; 51125 MW; 63B03903164CB1ED CRC64;

Alignment Scores:
Pred. No.: 44.2 Length: 458
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```


RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoco A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism *Streptomyces avermitilis*.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
DR EMBL; AP005035; BAC71139.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003760; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000524; HTH_GNTR.
DR InterPro; IPR000408; Reg_chr_condens.
DR Pfam; PF00392; GntR; 1.
DR PRINTS; PR00035; HTHGNTR.
DR SMART; SM00345; HTH_GNTR; 1.
DR PROSITE; PS50949; HTH_GNTR; 1.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KM Complete proteome; DNA-binding; Transcription;
KM Transcription regulation.
SQ SEQUENCE 478 AA; 50879 MW; 917EB2C508E3496E CRC64;

Alignment Scores:
Pred. No.: 44 Length: 478
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: Gaps: 0

US-10-030-271-1 (1-909) x Q82HT1 (1-478)

QY 838 CTGACTGAGCGCCGCGAGAGCTGTG 864

DB 25 LeuThrglualaleuArgglualalaval 33

RESULT 43

Q6DDP9 PRELIMINARY; PRT; 516 AA.

AC Q6DDP9; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
OS *Xenopus laevis* (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; *Xenopus*.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for *Xenopus* research: The NIH *Xenopus*
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshitoki S., Carrinci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwen J.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kesteman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywnski M.I., Skalka U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocytes;
RA Klein S., Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC077609; AA77609.1; -.
DR InterPro; IPR001849; PH.
DR InterPro; IPR01036; PH-related.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KM Hypothetical protein.
SQ SEQUENCE 516 AA; 56918 MW; DAF29FCE4B632580 CRC64;

Alignment Scores:
Pred. No.: 43.6 Length: 516
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: Gaps: 0

US-10-030-271-1 (1-909) x Q6DDP9 (1-516)

QY 464 AGCGGACGCGCGGAGTCCGCGCGC 490

DB 3 SerGlyserGlyGlyValGlyAlaGly 11

RESULT 44

Q6K663 PRELIMINARY; PRT; 525 AA.

AC Q6K663; 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative cytochrome P450.
GN Name=Q01003_F04.8;
OS *Oryza sativa* (Japanese cultivated rice).
OC Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; *Oryza*.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
DR EMBL; AP005285; BAD23217.1; -.
DR GO; GO:0004497; F:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR002401; BP4501.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00463; EP4501.
DR PROSITE; PS00385; P450.
KM Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 525 AA; 58605 MW; B1C48A46716C9283 CRC64;

Alignment Scores:
Pred. No.: 43.5 Length: 525
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.97% Indels: 0
 DB: Gaps: 0

US-10-030-271-1 (1-909) x Q6K663 (1-525)

QY 669 GCGGCTGCGCGCGAGCTGACGTGTT 695
 DB 171 G1YAlaG1YAla1a1a1aG1YArgVal 179

RESULT 45

Q75EB6 PRELIMINARY; PRT; 571 AA.

AC Q75EB6;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE AAR158WP.
 GN ORFNames=AAR158W;
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Eremothecium.
 NC NCBI_TaxId=33169;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10895;
 RA Dietrich F.S., Voegelé S.E., Brachat S., Ierch A., Gaffney T.,
 RA Philippson P.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016814; AAS50525.1; -.
 DR AAD; AAR158W; -.
 SQ SEQUENCE 571 AA; 59251 MW; 2CB758751F77D725 CRC64;

Alignment Scores:

Pred. No.: 43.1 Length: 571
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: Gaps: 0

US-10-030-271-1 (1-909) x Q75EB6 (1-571)

QY 277 GCACGCGCAGAGTTGCCCGACGACC 251
 DB 146 AlarGAlaG1YAla1a1aProAla1a 154

RESULT 46

Q871T5 PRELIMINARY; PRT; 629 AA.

AC Q871T5;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Hypothetical protein 7F4.020.
 GN Name=7F4.020;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 NC NCBI_TaxId=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohenseil J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX294020; CAD70906.1; -.
 DR InterPro; IPR006118; Recombinase_S.
 DR PROSITE; PS00398; RECOMBINASES_2; UNKNOWN_1.

KW Hypothetical protein.
 SQ SEQUENCE 629 AA; 67050 MW; 9A32B67A823466B0 CRC64;

Alignment Scores:

Pred. No.: 42.6 Length: 629
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.97% Indels: 0
 DB: Gaps: 0

US-10-030-271-1 (1-909) x Q871T5 (1-629)

QY 496 GGTGCTCCAGACGCGCGCGAGAGCG 522
 DB 350 G1YG1YAlaArgArgArgArgArgGly 358

RESULT 47

Q6XP25 PRELIMINARY; PRT; 636 AA.

AC Q6XP25;
 DT 05-JUL-2004 (TREMBLrel. 27, Created)
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
 DE Cyclophilin-like protein (EC 5.2.1.8).
 GN Name=Cyp-70;
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 NC NCBI_TaxId=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endosperm;
 RA Johnson J.C., Clarke B.C., Bhawe M.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PPIases accelerate the folding of proteins (By similarity).
 CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline (omega=0).
 CC -1- SIMILARITY: Belongs to the cyclophilin-type PPIase family.
 DR EMBL; AY17752; AAP44536.1; -.
 DR HSSP; Q27450; 1A33.
 DR GO; GO:0016853; F:isomerase activity; IEA.
 DR GO; GO:0003755; F:peptidyl-prolyl cis-trans isomerase activity; IEA.
 DR GO; GO:0006457; P:protein folding; IEA.
 DR InterPro; IPR002130; CSA_PPIase.
 DR Pfam; PF00160; Pro_isomerase; 1.
 DR PRINTS; PR00153; CSAPPISMRASE.
 DR PROSITE; PS50072; CSA_PPIASE_2; 1.
 DR KW isomerase; Rotamase.
 SQ SEQUENCE 636 AA; 69858 MW; B8E9DE395EB8FDB5 CRC64;

Alignment Scores:

Pred. No.: 42.5 Length: 636
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: Gaps: 0

US-10-030-271-1 (1-909) x Q6XP25 (1-636)

QY 224 CGGCGCTCAGCTCCAGACGAGCTCT 198
 DB 616 ArgArgSerSerSerArgSerSer 624

RESULT 48

Q8B2N7 PRELIMINARY; PRT; 709 AA.

AC Q8B2N7;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TRENBLrel. 26. Last annotation update)
 DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
 DE library, clone:933051F09 product:similar to THYROID HORMONE RECEPTOR-
 DE ASSOCIATED PROTEIN COMPLEX COMPONENT TRAP150.
 GN Name=Thrap3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RP STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Mech. Enzymol. 303:19-44 (1999).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Mech. Enzymol. 303:19-44 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuana T., Teshiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada K.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaishiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Maruyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tomaru A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK034044; BAC28557.1; --
 DR MGI; MGI:2442637; Thrap3.
 DR GO; GO:0000119; C:mediator complex; IDA.
 DR GO; GO:0006366; P:transcription from Pol II promoter; IC.

KW Receptor.
 SQ SEQUENCE 709 AA; 80110 MW; E573176A0648D7EC CRC64;
 Alignment Scores:
 Pred. No.: 42 Length: 709
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.00% Indels: 0
 DB: 2 Gaps: 0
 US-10-030-271-1 (1-909) x Q8B2N7 (1-709)
 Qy 224 CGGCGCTCCAGCTCCAGAGGCTCT 198
 Db 155 ArgArgSerSerSerSerArgSerSer 163
 RESULT 49
 ID 1150 HUMAN STANDARD; PRT; 955 AA.
 AC 09Y2M1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Thyroid hormone receptor-associated protein complex 150 kDa component
 DE (Thrap150) (Thyroid hormone receptor associated protein 3).
 GN Name=THRAP3; Synonyms=TRAP150;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 ON NCBI_TaxId=9606;
 RX MEDLINE=99214851; PubMed=10198638; DOI=10.1016/S1097-2765(00)80463-3;
 RA Ito M., Yuan C.-X., Malik S., Gu W., Fondell J.D., Yamamura S.,
 RA Fu Z.-Y., Zhang X., Qin J., Roeder R.G.;
 RT "Identity between TRAP and SMC complexes indicates novel pathways for
 RT the function of nuclear receptors and diverse mammalian activators.";
 RL Mol. Cell 3:361-370(1999).
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Diencephalon;
 RX MEDLINE=99214851; PubMed=10198638; DOI=10.1016/S1097-2765(00)80463-3;
 RA Ito M., Yuan C.-X., Malik S., Gu W., Fondell J.D., Yamamura S.,
 RA Fu Z.-Y., Zhang X., Qin J., Roeder R.G.;
 RT "Identity between TRAP and SMC complexes indicates novel pathways for
 RT the function of nuclear receptors and diverse mammalian activators.";
 RL Mol. Cell 3:361-370(1999).
 RN [2]
 RP SEQUENCE OF 1-672 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Uedon T.B., Tohiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Iqbaliano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunnarson P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Botterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schenck A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
 CC -1- FUNCTION: Plays a role in transcriptional coactivation.
 CC -1- SUBUNIT: Subunit of the large multiprotein complex TRAP.
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF117756; AAD22034.1; -

DR EMBL; BC137554; AAI37554.1; -

DR Genew; HGNC:22964; THRAP3.

DR MIM; 603809; -

DR GO; GO:0000119; C:mediator complex; IDA.

DR GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0030374; F:ligand-dependent nuclear receptor transcrip. .; NAS.

DR GO; GO:0004872; F:receptor activity; IDA.

DR GO; GO:0016455; F:RNA polymerase II transcription mediator ac. .; IDA.

DR GO; GO:0046966; F:thyroid hormone receptor binding; IDA.

DR GO; GO:0016583; F:transcriptional activator binding; NAS.

DR GO; GO:0042809; F:vitamin D receptor binding; IDA.

DR GO; GO:0030521; P:androgen receptor signaling pathway; IDA.

DR GO; GO:0006367; P:transcription initiation from Pol II promoter; IDA.

KW Activator; ATP-binding; Nuclear protein; Receptor;

KW Transcription regulation.

KW NP_BIND 552 559 ATP (Potential).

FT DOMAIN 7 339 Ser-rich.

FT DOMAIN 12 161 Arg-rich.

FT CONFLICT 669 669 N -> K (in Ref. 2).

FT CONFLICT 672 672 S -> K (in Ref. 2).

SQ SEQUENCE 955 AA; 106893 MW; FF031C3632E1B6A8 CRC64;

Alignment Scores:

Pred. No.: 40.5 Length: 955

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.00% Indels: 0

DB: 1 Gaps: 0

US-10-030-271-1 (1-909) x T150_HUMAN (1-955)

QY 224 CGCGCTCCAGCTCCAGCAGAGCTCT 198

Db 155 ArgArgSerSerSerSerArgSerSer 163

RESULT 50

Q7XM40 PRELIMINARY; PRT; 1099 AA.

AC Q7XM40; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE OSJNB0022P19.2 protein.

GN Name=OSJNB0022P19.2;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Erihartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=12447439; DOI=10.1038/nature01183;

RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J., Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y., Weng Q., Zhang L., Lu Y., Mu Y., Sun T., Lei H., Li T., Hu H., Guan J., Liu X., Lu T., Li C., Wu Y., Sun T., Lu Y., Zhang L.S., Yu Z., Pan D., Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H., Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y., Chen J., Chen H., Shao C., Sun Y., Hu Q., Zhang X., Zhang W., Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W., Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y., Han B.;

RT "Sequence and analysis of rice chromosome 4.";

RL Nature 420:316-320(2002).

DR EMBL; AL663001; CAE04807.2; -

DR Gramene; Q7XM40; -

DR GO; GO:0003677; F:DNA binding; IEA.

DR GO; GO:0006310; F:DNA recombination; IEA.

DR InterPro; IPR001584; Rve

DR InterPro; IPR001878; Znf_CCHC.

DR Pfam; PF00665; Ives; 1.

DR Pfam; PF00098; Zf_CCHC; 1.

DR PRINTS; PR00939; C2HCZNFINGER.

DR SMART; SM00343; ZNF_C2HC; 1.

DR PROSITE; PS50158; ZF_CCHC; 1.

SQ SEQUENCE 1099 AA; 121665 MW; 35C5BBA146ED5A86 CRC64;

Alignment Scores:

Pred. No.: 39.8 Length: 1099

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.97% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-1 (1-909) x Q7XM40 (1-1099)

QY 663 GCGCCAGCGCTGCGCGCGAGCTGGA 689

Db 3 AlaproglyAlaGlyAlaAlaAlaGly 11

Search completed: March 22, 2005, 03:20:18

Job time: 176.229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nzp model

Run on: March 22, 2005, 01:55:42 ; Search time 126.322 Seconds
(without alignments)
5566.160 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 303
Sequence: 1 atggcgtatccgggtcgcac.....tcagtgatgagcgtcgcac 909

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3967864

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

MODEL=frame_nzp.model -DBV=xlp
-O=/cgn2.1/USPTO_spool_p/US10030271/runat_21032005_153546_26816/app_query.fasta_1.3150
-DB=A.GeneSeq_16Dec04 -QFMT=fastlan -SUFFIX=oligo_rag -MINMATCH=40.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10030271.@CNC_1.1.445.@runat_21032005_153546_26816 -NCPU=6 -ICPU=3
-NO MAP -LARGEOBURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: GeneSeq_16Dec04:*
2: GeneSeq_19808:*
3: GeneSeq_19908:*
4: GeneSeq_20008:*
5: GeneSeq_20028:*
6: GeneSeq_20038:*
7: GeneSeq_20038:*
8: GeneSeq_20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	303	100.0	303	4	AAB60386 Human apo
2	303	100.0	326	4	AAB60387 Human apo
3	303	73.9	304	3	AAB15551 ApocB8
4	224	73.9	318	5	AAB24860 Human DED
5	224	73.9	318	5	AAB38903 Human DED
6	224	73.9	326	5	AAB807263 Human apo
7	224	73.9	326	7	ADCT9280 Human DED
8	224	73.9	366	4	AAG73852 Human col
9	194	64.0	319	3	AAB58420 Lung can
10	134	44.2	242	4	AAM25705 Human pro

11	120	39.6	361	5	ABB06038 Human NS
12	120	39.6	368	5	ABB06039 Human NS
13	67	22.1	101	5	AAE24855 Human DED
14	67	22.1	101	5	AAE38898 Human DED
15	66	21.8	217	4	AAM41591 Human pol
16	66	21.8	277	4	AAM39805 Human pol
17	13	4.3	318	2	AAW90108 Human FLA
18	13	4.3	318	2	AAW90109 Mouse FLA
19	13	4.3	318	3	AAV51023 Murine DE
20	13	4.3	318	3	AAV51022 Human DED
21	13	4.3	318	4	AAB93016 Human pro
22	13	4.3	318	4	AAB94040 Human pro
23	13	4.3	318	5	AAE26087 Human FLA
24	13	4.3	318	5	AAE26088 Mouse FLA
25	10	3.3	104	6	ABM38212 Pseudonib
26	10	3.3	104	6	ABM38212 Pseudonib
27	10	3.3	314	4	AAU30377 Novel hum
28	9	3.0	51	4	AAU55376 Pseudonib
29	9	3.0	51	4	AAU41040 Pseudonib
30	9	3.0	51	6	ABM37559 Pseudonib
31	9	3.0	51	6	ABM51895 Pseudonib
32	9	3.0	68	8	ADN01127 Human cel
33	9	3.0	71	4	AAU41429 Pseudonib
34	9	3.0	71	6	ABM37948 Pseudonib
35	9	3.0	89	4	AAU62949 Pseudonib
36	9	3.0	89	6	ABM59468 Pseudonib
37	9	3.0	91	7	ABO81797 Pseudonib
38	9	3.0	125	4	AAU46695 Pseudonib
39	9	3.0	125	6	ABM43214 Human gen
40	9	3.0	211	4	ABG63622 Human alb
41	9	3.0	211	5	ADG57327 Human sec
42	9	3.0	211	8	ADL76887 Albunin f
43	9	3.0	211	8	ABP62772 Streptomy
44	9	3.0	262	7	ADJ72183 Pseudonib
45	9	3.0	262	7	ABO69707 Human DIT
46	9	3.0	273	6	ABR41337 Coryneb
47	9	3.0	331	7	ABO69778 Coryneb
48	9	3.0	331	7	ABO69737 Coryneb
49	9	3.0	336	4	AAE79487 Pseudonib
50	9	3.0	336	4	ABO70227 Pseudonib
51	9	3.0	350	7	ABO69707 Pseudonib
52	9	3.0	350	7	ABO69707 Pseudonib
53	9	3.0	369	8	ABM83418 Human int
54	9	3.0	370	2	AAV56108 Human int
55	9	3.0	370	3	AAI19786 Human Wnt
56	9	3.0	370	3	AAI57594 Human Wnt
57	9	3.0	370	3	AAI70733 Human Wnt
58	9	3.0	370	6	ABU55882 Human WNT
59	9	3.0	370	6	AAE34038 WNT-1 pro
60	9	3.0	370	8	ADO08151 Human Wnt
61	9	3.0	370	8	ADO22214 Human Wnt
62	9	3.0	370	8	ADO49095 Human ded
63	9	3.0	386	7	ABO76378 Pseudonib
64	9	3.0	395	4	AAE92890 C glutam1
65	9	3.0	397	8	ADN99473 Novel hum
66	9	3.0	456	8	ADG42453 Bacterial
67	9	3.0	503	3	AB14247 Ced-3 pro
68	9	3.0	504	7	ABO74628 Pseudonib
69	9	3.0	505	2	AAE47463 Ced-3.3/
70	9	3.0	565	4	AAU39201 Human pol
71	9	3.0	657	7	ABO79255 Pseudonib
72	9	3.0	955	3	AB42231 Human ORF
73	9	3.0	955	7	ADJ70166 Human hea
74	9	3.0	955	8	ADP55415 Human TRA
75	9	3.0	962	4	AB11976 Human TRA
76	9	3.0	2031	6	ABU34762 Protein e
77	8	2.6	10	4	AAE95334 Human com
78	8	2.6	10	4	AAE95334 Human com
79	8	2.6	11	6	AAO31170 Pion (Pr
80	8	2.6	15	2	AAW37662 Peppy mot
81	8	2.6	15	2	AAW38911 Peptide r
82	8	2.6	15	7	ADB49288 Novel WW
83	8	2.6	21	4	AAE45725 Human 7TM

C 84	8	2.7	50	4	Aau49795	Propionib
C 85	8	2.7	50	6	Abm46314	Propionib
C 86	8	2.6	52	2	Aay20499	Human neu
C 87	8	2.7	62	8	ABO58857	Human gen
C 88	8	2.6	65	7	ABO62516	Klebsiell
C 89	8	2.6	73	4	ABG00429	Novel hum
C 90	8	2.7	73	4	ABG13593	Novel hum
C 91	8	2.7	73	6	ABO00821	Polypepti
C 92	8	2.6	76	3	AAg44964	Zea mays
C 93	8	2.6	76	7	ABO63430	Zea mays
C 94	8	2.7	79	5	AAU69709	Cell deat
C 95	8	2.7	80	8	ADP03245	S kanamyc
C 96	8	2.7	91	3	AAg22746	Zea mays
C 97	8	2.7	92	3	AAg32826	Zea mays
C 98	8	2.7	95	4	AAU44719	Propionib
C 99	8	2.6	95	4	AAU54277	Propionib
C 100	8	2.7	95	6	ABW41238	Propionib

ALIGNMENTS

RESULT 1

ID AAB60386 standard: protein; 303 AA.

AC AAB60386;

DT 24-APR-2001 (first entry)

DE Human apoptosis-associated factor NT2RM1000558 (portion), SEQ ID NO:2.

KW Human; apoptosis-associated factor; NT2RM1000558; death effector domain;

KM DEB; caspase family cleavage domain; pro-apoptotic; drug screening;

KN cell proliferation; ischaemic disease; chronic viral disease.

OS Homo sapiens.

XX MO200104300-A1.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000MO-JP004516.

XX 08-JUL-1999; 99JP-00194179.

XX 18-OCT-1999; 99US-0159586P.

XX (HELI-) HELIX RES INST.

XX Ota T, Isegai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;

XX WPI; 2001-138348/14.

XX N-PSDB; AAF27407.

XX Polynucleotide encoding an apoptosis-associated factor protein with death

XX effector domain and caspase family-cleavage domain, useful in regulating

XX diseases with cell proliferation.

XX Claim 1; Page 43-44; 53pp; Japanese.

XX The invention relates to a novel human apoptosis-associated factor

XX (AAB60386, AAB60387), designated NT2RM1000558, which contains a death

XX effector domain (DED) and a caspase family cleavage domain and is capable

XX of inducing apoptosis in cells. The invention also relates to nucleic

XX acids encoding the protein (AAF27407, AAF27408); variants of the protein

XX (particularly dominant negative variants); vectors and host cells

XX comprising a nucleic acid which encodes an apoptosis-associated factor

XX of the invention; the recombinant production of the protein; an antibody

XX against the protein; and methods of screening for compounds which can

XX regulate apoptosis. The apoptosis-related factor is useful in regulating

XX diseases associated with cell proliferation and in screening drug

XX candidates e.g., for regulating cell proliferation or cell death in

XX ischaemic diseases and chronic viral diseases. The present sequence

XX represents a substantial proportion of the human apoptosis-associated

CC	factor NT2RM1000558
XX	Sequence 303 AA;
SQ	Alignment Scores:
	Pred. No.: 2.14e-266
	Score: 303.00
	Percent Similarity: 100.00%
	Best Local Similarity: 100.00%
	Query Match: 100.00%
	DB: 4
	Gaps: 0
US-10-030-271-1 (1-909) x AAB60386 (1-303)	
QY	1 ATGGGCTATCCGGGATCGACCCCGCCGCTGCTGGAGAGATGATGCTGAGTAC 60
DB	1 MetAlaLeuSerGlySerThrProAlaProCySerTrpGluAmpGluCysLeuAspPyr 20
QY	61 TACGGGATGCTGTGCTTACCCGTAATGTCAGGTGGGCGGCACTGACGAGTGC 120
DB	21 TyrGlyMetLeuSerLeuHisArgMetPheGluValValGlyGlyGlnLeuThrGluCys 40
QY	121 GAGCTGAGAGCTCCGGGCTTTCTGCTGATGAGGCTCCTGGCGCGCGAGGCTTACCC 180
DB	41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60
QY	181 CGGGCCCGACCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGGCGGAG 240
DB	61 ArgAlaArgSerGlyLeuGluLeuLeuLeuGluLeuGluValArgGlyGlnCysGlyVal 80
QY	241 AGCAACTGCGGCTGCTGGGCAACTCTGCGGCTGCTGGCGCGCGACGACTGTCGC 300
DB	81 SerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeuPro 100
QY	301 CACCTGGCGGCAAGCGGCGCGCGAGTGTCTCAGAGAGCTATGAGGACCTCC 360
DB	101 HisLeuAlaArgLysArgArgArgProValSerProGluArgLysSerTrpGlyThrSer 120
QY	361 AGCTCTTCAAGAGAGACAGAGGATAGCTGCGCTGCGGCAAGTCAAGATTTTCA 420
DB	121 SerSerSerLysArgThrGlnGlySerCysArgArgArgGlnSerSerSerSerAla 140
QY	421 AATTCTACAGCGGCTCACTGGAGACAGGCTCCCCCAACCAAGCGGCGAGGCGAGT 480
DB	141 AsnSerGlnGlnGlnGlnTrpGlnThrGlySerProPorphLysArgGlnArgArgSer 160
QY	481 CGGGGCGCGGCGGAGTGGTGGCCAGACGCGGCGGAGAGGCGGCGGCGGCGGAG 540
DB	161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgArgGlyAlaProAlaAlaProGln 180
QY	541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAGAGCAAGTGAACCTGTGATCCGGCTC 600
DB	181 GlnGlnSerGluProAlaArgProSerSerGlnGlyValThrCysAspIleArgGlu 200
QY	601 CGGGTTCAGACAGTACTGCGAGCATGGCCAGCTTGGACGAGGCGGTGGCATCCCGG 660
DB	201 ArgValArgAlaGluArgCysGluHisGlyProAlaLeuGlnGlnGlyValAlaSerArg 220
QY	661 CGGCCCCAGCGGCTGGGCGGCGGCGGAGTGGAGTGGGCGAGGCGGCGGCTGCTGC 720
DB	221 ArgProGlnAlaLeuAlaArgGlnLeuAspAlaPheGlyGlnAlaAlaValaValaLeuArg 240
QY	721 TCAAGGAGACTGGGCTGTGTGTTGTGACATCAAGTTCTCAGAGCTTCTATCTGGAC 780
DB	241 SerArgAspLeuGlySerValValCysAspIleLysPheSerArgLysLeuSerTrpLeuAsp 260
QY	781 GCTTTTGGGCGCACTACCTGAGTGGCGCTGCTGACAGGCGCTGGGCGGCTGTCTCT 840
DB	261 AlaPheTrpGlyAspTrpLysSerGlyAlaLeuLeuGlnAlaLeuArgGlyValaPheLeu 280
QY	841 ACTAGAGCGCGCAGAGGCGGCTGGCGGAGGCTGTGGCTGGCTGGCTGAGTGGAT 900
DB	281 ThrGluAlaLeuArgGluAlaValaGlyArgGluAlaValaArgGluLeuValaSerValaAsp 300

QY 901 GAGGCTGAC 909
 |||||
 Db 301 GluAlaAsp 303

RESULT 2
 AAB60387
 ID AAB60387 standard; protein; 326 AA.

XX AAB60387;

XX 24-APR-2001 (first entry)

XX Human apoptosis-associated factor NT2RM1000558, SEQ ID NO.4.

XX Human; apoptosis-associated factor; NT2RM1000558; death effector domain;

XX DEB; caspase family cleavage domain; pro-apoptotic; drug screening;

XX KW cell proliferation; ischaemic disease; chronic viral disease.

XX OS Homo sapiens;

XX MO200104300-A1.

XX 18-JAN-2001.

XX 06-JUL-2000; 2000WO-JP004516.

XX 08-JUL-1999; 99JP-00194179.

XX 18-OCT-1999; 99US-0159586P.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;

XX WPI; 2001-138348/14.

XX N-PSDB; AAF27408.

XX polynucleotide encoding an apoptosis-associated factor protein with death

XX PT effector domain and caspase family-cleavage domain, useful in regulating

XX PT diseases with cell proliferation.

XX PS Claim 3; Page 47-48; 53pp; Japanese.

XX The invention relates to a novel human apoptosis-associated factor

XX (AAB60386, AAB60387), designated NT2RM1000558, which contains a death

XX effector domain (DED) and a caspase family cleavage domain and is capable

XX of inducing apoptosis in cells. The invention also relates to nucleic

XX acids encoding the protein (AAE27407, AAF27408); variants of the protein

XX (particularly dominant negative variants); vectors and host cells

XX comprising a nucleic acid which encodes an apoptosis-associated factor

XX against the protein; and methods of screening for compounds which can

XX regulate apoptosis. The apoptosis-related factor is useful in regulating

XX diseases associated with cell proliferation and in screening drug

XX candidates e.g., for regulating cell proliferation or cell death in

XX CC ischaemic diseases and chronic viral diseases. The present sequence

XX CC represents the human apoptosis-associated factor NT2RM1000558

XX SQ Sequence 326 AA;

Alignment Scores:

Pred. No.: 2,11e-266

Score: 303.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 100.00%

DB: 4

US-10-030-271-1 (1-909) x AAB60387 (1-326)

Length: 326

Matches: 303

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

1 AATGGCGCTATCCGGGTGACACCCGCCGCTGCGAGAGAGATAGTGCCTGAGTAC 60

1 MetAlaLeuSerGlySerThrProAlaProCystrGluGluAspGluCysLeuAspTyr 20

QY 61 TACGGAGATGCTGTCCGCTTACCCGTAATGTTTCAGAGTGTGGCGGCAACTGACGAGTGC 120
 |||||
 Db 21 TyrGlyMetLeuSerLeuHisArgMetPheGluValValGlyGlyLeuThrGluCys 40

QY 121 GAGCTGAGAGCTCTGCGCTTTCTGCTGATGAGGCTCTCTGCGCGCGGAGGCTTATGCC 180
 |||||
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyValAlaGlyGlyLeuAla 60

QY 181 CGGGCCCCGAGCGGCTGAGAGCTCTGCTGAGAGCTGAGAGCGCGGCGAGTGGCGGAG 240
 |||||
 Db 61 ArgAlaArgSerGlyLeuGluLeuLeuLeuGluLeuGluValArgGlyGlyCysGlyGlu 80

QY 241 AGCAACTGCGGCTCTGCGGCAACTCTGCGCGTGTGAGCGCGGCGGAGGCTGTCGCG 300
 |||||
 Db 81 SerAlaLeuArgLeuLeuGluGlyGluLeuLeuAlaValAlaAlaArgHisAspLeuLeuPro 100

QY 301 CACTGGCGCGGAGCGCGCGCGCGAGTGTCTCCAGAACGCTATAGTATGACCTTCC 360
 |||||
 Db 101 HisLeuAlaArgIleArgIleArgIleArgIleArgIleArgIleArgIleArgIleArgIle 120

QY 361 AGCTCTTCAAG 420
 |||||
 Db 121 SerSerSerLeuArgIleArgIleArgIleArgIleArgIleArgIleArgIleArgIleArgIle 140

QY 421 AATTCACAG 480
 |||||
 Db 141 AsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160

QY 481 CGGGGCGCGCGCGAGTGTGTGTCAGACCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 540
 |||||
 Db 161 ArgGlyArgProSerGlyGlyValAlaArgIleArgIleArgIleArgIleArgIleArgIleArgIle 180

QY 541 CAGCAGTCAAG 600
 |||||
 Db 181 GlnGlnSerGlnProAlaArgProSerSerGlnGlyGlyValThrCysAspIleArgLeu 200

QY 601 CGGGTTCGAG 660
 |||||
 Db 201 ArgValAlaArgIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220

QY 661 CGGCCCCAGCGCGCTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 |||||
 Db 221 ArgProGlnAlaLeuAlaArgIleLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240

QY 721 TCAAGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 |||||
 Db 241 SerArgAspLeuGlySerValAlaCysAspIleLeuPheSerGluLeuSerLeuAsp 260

QY 781 GCCTTCTGAGGAG 840
 |||||
 Db 261 AlaPheThrGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyAlaPheLeu 280

QY 841 ACTGAGCGCTGCGAG 900
 |||||
 Db 281 ThrGluAlaLeuArgGluValAlaGlyArgGluAlaValAlaArgLeuLeuValSerValAsp 300

QY 901 GAGGCTGAC 909
 |||||
 Db 301 GluAlaAsp 303

RESULT 3
 AAB15551
 ID AAB15551 standard; protein; 304 AA.

XX AAB15551;

XX 28-FEB-2001 (first entry)

XX Apoptosis related protein encoded by gene 1 clone HLDOK36.

XX Cytostatic; anti-inflammatory; immunosuppressive; anti-ischaemic; cardiant;

XX KW virocidal; anti-AIDS; vasotrophic; anti-ischaemic; antiparkinsonian;

XX	anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;
KW	colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;
KW	autoimmune disorder; multiple sclerosis; viral infection.
OS	Homo sapiens.
XX	
XX	MO200056752-A2.
PN	
PD	28-SEP-2000.
XX	
PF	15-MAR-2000, 2000MO-US006642.
XX	
PR	24-MAR-1999; 99US-0126018P.
PR	17-JUN-1999; 99US-0139638P.
PR	18-AUG-1999; 99US-0149449P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Ni J, Young PA;
XX	
DR	WPI; 2000-587660/55.
DR	N-PSDB; AAA95790.
XX	
PT	Nucleic acids encoding human apoptosis associated protein, useful for the
PT	prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's
PT	disease, inflammation and ischemic injury.
XX	
PS	Claim 11, Page 252-253; 273pp; English.
XX	
CC	The invention relates to the isolation of genes encoding 9 human
CC	apoptosis-related proteins. The nucleotide sequences AAA95790-A95798
CC	encode the human apoptosis related proteins AAB15551-B15559. The genes
CC	can be used to generate fusion proteins by linking to the gene for the
CC	human immunoglobulin G Fc (IgG Fc) portion (AAA95799) for increasing the
CC	stability of the fusion protein as compared to the human protein only.
CC	The gene and encoded protein may be used in the prevention, treatment and
CC	diagnosis of diseases associated with inappropriate apoptosis associated
CC	protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,
CC	melanomas and lymphomas), inflammation, autoimmune disorders (e.g.
CC	multiple sclerosis) and viral infections (e.g. herpes)
XX	
SO	Sequence 304 AA;
Alignment Scores:	
Pred. No.:	1,51E-194
Score:	224.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	73.93%
DB:	3
	Gaps: 0
US-10-030-271-1 (1-909) x AAB15551 (1-304)	
QY	238 GAGAGCAACCTGCGGCTGCTGGGAGCAACTCTCGCGTGTGACCGCAGCACTGCTG 297
Db	58 GlnSerAsnIleuAlygrIuAlygrAlygrAlygrAlygrAlygrAlygrAlygrAlygrAlygr 77
QY	298 CCGCACCCTGGCGCGCAAGCGCGCGCGCGCAGTGTCTTCAGAAACGCTATAGTATGCACC 357
Db	78 ProHisIleuAlyAlygrAlygrAlygrAlygrAlygrAlygrAlygrAlygrAlygrAlygr 97
QY	358 TTCAGCTCTTCAAAGGAGCAGAGGGTGTGTCGCTGCGCGCGAGTAAAGATTC 417
Db	98 SerSerSerSerIleAlygrIuAlygrIuAlygrIuAlygrIuAlygrIuAlygrIuAlygrIu 117
QY	418 GCAAAATTCACAGCAGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGCAGCGCG 477
Db	118 AlaAsnSerGlnGlnIuAlygrIuAlygrIuAlygrIuAlygrIuAlygrIuAlygrIuAlygr 137
QY	478 AGTCGGGGCGCGCGCAGTGGTGTGTCACAGCGCGCGCGAAGGGGCGCCAGCGCGCACCC 537
Db	138 SerArgGlyArgProSerGlyGlyIuAlygrAlygrAlygrAlygrAlygrAlygrAlygrAlygr 157

QY	538	CAGCAGCACTGAGAGCCGCCAGACTCTTCTGTGAAGCAAAAGTGAACCTGTGACATCCG	597		
Db	158	GInGInGInSeGclPrOAlaIRProSeSerSeRtUgLySeValTmTySaMpIlLeArg	177		
QY	598	CTCCGGGTTTGAGAGCAAGTACTGCGAGCATGSGCCAGCCTTTGAGAGCGGCGTGGCATCC	657		
Db	178	LeuArgValaIRaGlaGluTyrcYseGluHfISglProAlaLeuGluGInGlyValAlaSer	197		
QY	658	CGCGCGGCCCCAGGCGCTGGCGCGAGCTGSAcGTGTTTGSGGAGGCGCACCGAGTGTG	717		
Db	198	ArgArgPrIGlnAlaLeuAlaArgGlnLeuSuSpAlPheGlyGlnAlaTrAlaValLeu	217		
QY	718	CGCTCAAGGAGCCTGGGCTGTGTGGTTTGACATCAAGTTCTTCAGAGCTTCTCTATCTG	777		
Db	218	ArgSerArgaSpLeuGlySerValValCysaPheIleUyPheSerIleuLeuSerTyLeu	237		
QY	778	GACGCTTTTGSGGGGAGTACTCTGATGSGGCCCTTGCTGCAGAGGCCCTGGCGGCGTGTTC	837		
Db	238	AspAlaPheTrpGlyaSpTyIleuSerIyAlaLeuLeuGlnAlaLeuArgGlyValPhe	257		
QY	838	CTGACTGAGGCGCCTCGAGAGGGCTGTGGGCGGAGAGCGTGTTCGCTGTGTGTCAGTGTG	897		
Db	258	LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValaArgLeuLeuValSerVal	277		
QY	898	GATGAGCTGTAC 909			
Db	278	AspGlnAlaAsp 281			
RESULT 4					
ID	AAE24860	standard; protein; 318 AA.			
AC	AAE24860;				
XX					
DT	22-OCT-2002	(first entry)			
XX					
DE	Human DED4 (death effector domain) protein.				
XX					
KW	NB; death domain; DD; death effector domain; DED; Chlamydia infection;				
KW	NB-AAC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;				
KW	inflammation; allergy; autoimmunity; allograft rejection; cell division;				
KW	immune-based pathology; fibrosis; arthritis; graft versus host disease;				
KW	immunosuppressive; gene therapy; antisense therapy.				
XX					
OS	Homo sapiens.				
XX					
XX	Key	Location/Qualifiers			
FT	Region	53..74			
FT		/note="Nuclear localisation sequence"			
XX					
XX	MO200240680-A2.				
XX					
XX	23-MAY-2002.				
XX					
XX	15-NOV-2001; 2001WO-US044844.				
XX					
PR	17-NOV-2000; 2000US-00715893.				
PR	29-JUN-2001; 2001US-0301889P.				
XX					
XX	(BURN-) BURNHAM INST.				
XX					
PI	Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;				
PI	Stemmer-Liawen F;				
XX					
DR	WPI; 2002-500222/53.				
XX					
XX	N-PSDB; AAD40080.				
XX					
PT	New polypeptide comprising a death domain or death effector domain,				
PT	useful for discovery of drugs that suppress infection, inflammation,				
PT	allergy, sepsis, autoimmunity, allograft rejection and other diseases.				
XX					
XX	Claim 3; Page 186-187; 209pp; English.				
XX					

CC provides death domain containing protein such as Chlamydia trachomatis death domain containing protein (CTDD) DD and neural growth factor receptor-interacting death domain (NIDD) DD. The invention is useful for identifying a binding agent (e.g. protein or drug) that binds a DD, DED or NB-ARC domain from DAP3, IRAK4, CTDD, DED or NIDD with a candidate binding agent and identifying an effective agent (e.g. protein or drug) that modulates the association of a DD, DED or NB-ARC domain with protein that binds the DD, DED or NB-ARC domain. The invention is also useful for modulating the level of cell processes such as apoptosis, cell adhesion, cell proliferation, cell stress responses, responses to microbial infection and B cell immunoglobulin class switching. DEDs, DEDs and NB-ARC domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are useful for discovery of drugs that suppress infection, autoimmunity, inflammation, allergy, allograft rejection, sepsis and other diseases. DD, DED or NB-ARC domain proteins are used to treat infection, allergy, autoimmunity, inflammation, allograft rejection, sepsis, keratinocytic hyperplasia, neoplasia, keloid, benign prostatic hyperplasia, fibrosis, inflammatory hyperplasia and smooth muscle cell proliferation in arteries following balloon angioplasty (restenosis). The invention is also used in antibody therapy and gene therapy. The present sequence is human DED4 protein. The DED4 gene is located on chromosome 19

CC
XX
SQ Sequence 318 AA;

Alignment Scores:
Pred. No.: 1.5e-194 Length: 318
Score: 224.00 Matches: 224
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.93% Indels: 0
Gaps: 0

US-10-030-271-1 (1-909) x AAB38903 (1-318)

QY 238 GAGAGCAACCTGCGGCTGCGGAGCAACTCTGCGGCTGCGGCGGAGCACTGCTG 297
DB 80 GUSerAenLeuAArgLeuLeuGlyGlnLeuLeuAValLeuAAlaRHisAspLeuLeu 99
QY 298 CCGCACTGCGGCGGAGCGCGCGGCGGAGTGTCTCCAGAAAGCTATAGTGGACCC 357
DB 100 ProHsLeuAAlaRHisAspLeuLeuAValLeuAAlaRHisAspLeuLeu 119
QY 358 TCCAGCTCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 417
DB 120 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 139
QY 418 GCAAAATTCAGCAGAGGAGTGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 477
DB 140 AlaAenSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 159
QY 478 AGTCGGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 537
DB 160 SerAArgGlyArgProSerSerGlyGlyAlaArgAArgAArgAArgAArgAArgAArg 179
QY 538 CAGCAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 597
DB 180 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
QY 598 CTCCGGGCTTGCAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
DB 200 LeuAArgValArgAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 219
QY 658 CCGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 717
DB 220 ArgAArgProGlnAlaLeuAAlaArgGlnLeuAAspValPheGlyGlnAlaAlaValLeu 239
QY 718 CGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 777
DB 240 ArgSerAArgAAspLeuGlySerValValCysAspPheLeuPheSerGlnLeuSerLeu 259
QY 778 GAGCGCTTCTGCGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 837
DB 260 AspAlaPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 279

QY 838 CTGACTGAGGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
DB 280 LeuThrGlnAlaLeuAArgGlnAlaValGlyArgGlnAlaValAArgLeuLeuValSerVal 299
QY 898 GATGAGGCTGAC 909
DB 300 AspGlnAlaAsp 303

RESULT 6

AB07263

ID AB07263 standard; protein; 326 AA.

XX AB07263;

DT 26-MAR-2002 (first entry)

DE Human apoptosis regulator (APR) polypeptide (Incyte ID. 3102521CD1).

XX APR; apoptosis regulator; cytosolic; antiatherosclerotic; osteopathic;

KW antiatherosclerotic; hepatotropic; antipsoriatic; antihelminthic; human;

KW antiatherogenic; antiatherogenic; antiatherogenic; anti-HIV; cancer;

KW antiinflammatory; antidiabetic; antitumor; nephrotropic; ophthalmologic;

KW immunosuppressive; dermatologic; antitumor; antiparasitic; fungicide;

KW antitumor; antibacterial; virucide; antiparasitic; protozoicide;

KW tranquillizer; vulnery; gynecological; vasotropic; gene therapy.

XX Homo sapiens.

XX WO200192527-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US017581.

XX 01-JUN-2000; 2000US-0209407P.

XX 30-NOV-2000; 2000US-0250326P.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YF, Azimzai Y, Yue H, Burford N, Ding L, Elliott VS;

XX Patterson C, Baughn MR;

XX WPI. 2002-114350/15.

XX N-PSDB; ABA94362.

XX Novel human apoptosis regulator polypeptides and polynucleotides for

PT diagnosing, preventing, treating cell proliferative, immunological and

PT reproductive disorders and for identifying modulators of therapeutic use.

XX Claim 1; Page 97-98; 103pp; English.

PS The invention provides human apoptosis regulator (APR) polypeptides and

XX polynucleotides. The APR polypeptides, polynucleotides and modulators

CC are useful for diagnosis, treatment and prevention of cell proliferative,

CC immunological and reproductive disorders. The cell proliferative

CC disorders include cancers, actinic keratosis, arteriosclerosis,

CC atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, and

CC immunological disorders include acquired immunodeficiency syndrome

CC (AIDS), adult respiratory distress syndrome, Addison's disease,

CC ankylosing spondylitis, amyloidosis, allergies, anemia, osteoporosis,

CC autoimmune hemolytic anemia, asthma, autoimmune thyroiditis, Crohn's

CC disease, contact dermatitis, diabetes mellitus, gout, Graves' disease,

CC glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus

CC erythematosus, systemic sclerosis, ulcerative colitis, haemodialysis,

CC uveitis; viral, bacterial, fungal, parasitic, protozoal, helminthic

CC infections and trauma. Reproductive disorders include disorders of

CC prolactin production, infertility, endometriosis, polycystic ovary

CC syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,

CC disruptions of spermatogenesis, cancer of testis and prostate, impotence,

CC carcinoma of male breast and gynecomastia. The APR polynucleotides are

CC useful for creating knockin humanized animals or transgenic animals to

CC model human disease and to detect and quantify gene expression in

CC biopsied tissues in which expression of ABRG is correlated with disease.
 CC APRG, fragments of it and antibodies specific for ABRG are useful as
 CC elements on a microarray which is useful to monitor or measure protein-
 CC protein interactions, drug-target interactions and gene expression
 CC profiles. The present sequence represents a human ABRG polypeptide
 XX

XX Sequence 326 AA;

Alignment Scores:

Pred. No.:	1-5e-194	Length:	326
Score:	224.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.93%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-1 (1-909) x ABB07263 (1-326)

QY 238 GAGAGCAACCTGCGGCTGTGGGCAACTCTGCGGCTGTGGCCGCGCAAGCACTGCTG 297
 DB 80 GtUSeArAsnLeuArGtLeuEnGtInLeuLeuArGtValLeuAlaRghIsaPLeuLeu 99
 QY 298 CCGCACTTGGCGGCGAGCGCGCGCGCAAGTGTCTTCCAGAACGCTATAGCTATGGCAAC 357
 DB 100 ProHIsLeuAlaRghIsaRghArGArGArGArGArGArGArGArGArGArGArGArGArG 119
 QY 358 TCACGCTCTTCAAGAGAGACAGAGAGGAGTGTGCGCGCGCGAGTCAACAGCTTCT 417
 DB 120 Ser 139
 QY 418 GCAAAATTCACAGAGGAGTGTGGGAGAGACAGGCTCCCGCCCAACAGCGGCGCGCG 477
 DB 140 AlaAenSerGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 159
 QY 478 AGTGGGCG 537
 DB 160 SerArGtGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtAr 179
 QY 538 CAGCAGCAGTCAAGCG 597
 DB 180 GInGInGInSerGInUProAlaArGProSerSerGInGtLyValThrCyAspIleArG 199
 QY 598 CTCGGGCTTCAAGAGAGTGTGCGAGCATGTGGCGCGCGCGCGCGCGCGCGCGCGCG 657
 DB 200 LeuArGtValArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGt 219
 QY 658 CCG 717
 DB 220 ArGArGtProGInAlaLeuAlaArGtInLeuAspValIphedGtInAlaThrAlaValLeu 239
 QY 718 CGGTCAAGGGAGCTGTGGCTCTGTGAGTGTGATCAAGTTCACAGGCTCTCTATCTG 777
 DB 240 ArGSeArAsnLeuArGtLeuEnGtInLeuLeuArGtValLeuAlaRghIsaPLeuLeu 259
 QY 778 GAGCGCTTGTGGGCGCACTACCTGAGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
 DB 260 AspAlaPhaRtGrGtLyAspTyrLeuSerGtAlaLeuLeuGInAlaLeuArGtLyValPhe 279
 QY 838 CTGACTGAGGCGCTGCGAGAGGCTGTGGCGCGGAGGAGGCTGTGGCTGTGCTGAGTGTG 897
 DB 280 LeuThrGtInAlaLeuArGtValGtArGtValaValaValaValaValaValaValaVala 299
 QY 898 GATGAGGCTGAC 909
 DB 300 AspGtAlaAsp 303

RESULT 7

ADCT9260

ADCT9260 standard; protein; 326 AA.

ADCT9260; 01-JAN-2004 (first entry)

XX Human DEDD2 protein SEQ ID NO:2.

DE human; death effector domains containing DNA-binding protein;
 XX DED-containing DNA-binding protein; DEDD2; cell death; gene therapy;
 KW cytostatic; cancer; chronic myeloid leukaemia.
 XX

OS Homo sapiens.

XX MO2003054195-A1.

XX 03-JUL-2003.

XX 20-DEC-2002; 2002WC-UP013371.

XX 20-DEC-2001; 2001JP-00387854.

XX 18-JUL-2002; 2002UP-00209458.

XX (MORG) MORINAGA MILK IND CO LTD.

XX Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;

XX WPI, 2003-569246/53.

XX DR N-PSDB; ADCT9259.

XX PT DNA encoding cell death proteins for treatment of kidney, large intestine

XX PS Claim 1; Page 20-21; 26pp; Japanese.

XX The present sequence represents a human death effector domains (DED)
 CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that
 CC causes cell death. Also described: (1) primer and probe for investigation
 CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
 CC cytostatic activity. DEDD2 can be used in the diagnosis and treatment of
 CC cancers of the kidney, large intestine and prostate, and acute and
 CC chronic myeloid leukaemia.
 XX

XX Sequence 326 AA;

Alignment Scores:

Pred. No.:	1.5e-194	Length:	326
Score:	224.00	Matches:	224
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	73.93%	Indels:	0
DB:	7	Gaps:	0

US-10-030-271-1 (1-909) x ADCT9260 (1-326)

QY 238 GAGAGCAACCTGCGGCTGTGGGCAACTCTGCGGCTGTGGCCGCGCAAGCACTGCTG 297
 DB 80 GtUSeArAsnLeuArGtLeuEnGtInLeuLeuArGtValLeuAlaRghIsaPLeuLeu 99
 QY 298 CCGCACTTGGCGGCGAGCG 357
 DB 100 ProHIsLeuAlaRghIsaRghArGArGArGArGArGArGArGArGArGArGArGArGArG 119
 QY 358 TCACGCTCTTCAAGAGAGACAGAGGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 417
 DB 120 Ser 139
 QY 418 GCAAAATTCACAGAGGAGTGTGGGAGAGAGGCTCCCGCCCAACAGCGGCGCGCG 477
 DB 140 AlaAenSerGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGInGIn 159
 QY 478 AGTGGGCG 537
 DB 160 SerArGtGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGtArGt 179
 QY 538 CAGCAGCAGTCAAGCG 597
 DB 180 GInGInGInSerGInUProAlaArGProSerSerGInGtLyValThrCyAspIleArG 199

QY 598 CTCGGGTTTCAGACAGTACTGCGAGCATGGGACCCTTGAGACGGCGTGGCATCC 657
 CC CC
 Db 200 LeuArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGluGlnGlyValAlaSer 219
 CC CC
 QY 658 CGGCGGCCCCCAGGCGCTGGCGGCGGACGCTGTTGGGACGGCCACCGCATGGCTG 717
 CC CC
 Db 220 ArgArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeu 239
 CC CC
 QY 718 CGCTCAAGGAGCCTGGGCTCTGTGGTTGTGACATCAATCTCAGAGCTCTCTATCTG 777
 CC CC
 Db 240 ArgSerArgAspLeuGlySerValValCysAspIleLeuPheSerGlnLeuSerTyrLeu 259
 CC CC
 QY 778 GACGCTTTTGGGGGAGCTACTGAGTGGCGCCCTGCTGCAGGCGCTGGCGGCGTTC 837
 CC CC
 Db 260 AspAlaPheTyrGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPhe 279
 CC CC
 QY 838 CTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCGCTGCTGCTGCTGCTG 897
 CC CC
 Db 280 LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerVal 299
 CC CC
 QY 898 GATGAGCTGAC 909
 CC CC
 Db 300 AspGlnAlaAsp 303
 CC CC
 RESULT 8
 AAG73852
 ID AAG73852 standard; protein; 366 AA.
 XX AC AAG73852;
 XX DT 03-SEP-2001 (first entry)
 XX DE Human colon cancer antigen protein SEQ ID NO:4616.
 XX KM Human colon cancer antigen protein SEQ ID NO:4616.
 XX KN colorectal carcinoma.
 XX OS Homo sapiens.
 XX PN WO200122920-A2.
 XX PD 05-APR-2001.
 XX PF 28-SEP-2000; 2000WO-US026524.
 XX PR 29-SEP-1999; 99US-0157137P.
 XX PR 03-NOV-1999; 99US-0163280P.
 XX PA (HUMA-) HUMAN GENOME SCT INC.
 XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX DR WPI: 2001-235357/24.
 XX DR N-PSDB; AAH33283.
 XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 XX PT useful for preventing, diagnosing and/or treating colorectal cancers.
 XX PS Claim 11; Page 6413-6414; 9803pp; English.
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytosolic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patient's own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acid
 CC into a host cell and culturing the cell to express the proteins. N and P

CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922
 XX Sequence 366 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1,476-194 Length: 366
 Score: 224.00 Matches: 224
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 73.93% Indels: 0
 DB: 4 Gaps: 0
 US-10-030-271-1 (1-909) x AAG73852 (1-366)
 QY 228 GAGAGCAACCTGCGGCTCTGCGGCACTCTCGCGGCTGTGGCCCGCCAGACCTGCTG 297
 CC CC
 Db 120 GUSerAsnLeuArgLeuLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 139
 CC CC
 QY 298 CCGCACCTGGCGCGCAAGCGGCGCGGCGGAGGTCCTCCAGAACGCTATAGCATGGCAC 357
 CC CC
 Db 140 ProHisLeuAlaArgLeuArgArgArgProValSerProGlnArgTyrSerTyrGlyThr 159
 CC CC
 QY 358 TCACAGCTCTTCAAGAGAGACAGAGAGGCTGCTCCGCTCGGACATCAAGCTTCT 417
 CC CC
 Db 160 SerSerSerSerLeuArgThrGlnGlySerCysArgArgArgGlnSerSerSer 179
 CC CC
 QY 418 GCAATTCTCAGCAGGCTGAGTGGGAGACAGGCTTCCCCCACCACCAAGCGGAGCGG 477
 CC CC
 Db 180 AlaAsnSerGlnGlnGlnGlnThrPglThrGlySerProProThrLeuArgGlnArgArg 199
 CC CC
 QY 478 AGTGGGCGCGGCGCCAGTGGTGTGTCGACAGCGCGCGGAGAGGGGCCCGACCGCACCC 537
 CC CC
 Db 200 SerArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaPro 219
 CC CC
 QY 538 CAGCAGCATGACAGAGCCCGCCAGACCTTCTGTGAAGCAAAAGTACCTGTGACATCCG 597
 CC CC
 Db 220 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArg 239
 CC CC
 QY 598 CTCGGGTTTCAGACAGTACTGCGAGCATGGGCCACTTGGAGACGGCGCTGGCATCC 657
 CC CC
 Db 240 LeuArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGlnGlnGlyValAlaSer 259
 CC CC
 QY 658 CGGCGGCCCCCAGGCGCTGGCGGCGGACGCTGTTGGGACGGCCACCGCATGGCTG 717
 CC CC
 Db 260 ArgArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeu 279
 CC CC
 QY 718 CGCTCAAGGAGCCTGGGCTCTGTGGTTGTGACATCAATCTCAGAGCTCTCTATCTG 777
 CC CC
 Db 280 ArgSerArgAspLeuGlySerValValCysAspIleLeuPheSerGlnLeuSerTyrLeu 299
 CC CC
 QY 778 GACGCTTTTGGGGGAGCTACTGAGTGGCGCCCTGCTGCAGGCGCTGGCGGCGTTC 837
 CC CC
 Db 300 AspAlaPheTyrGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPhe 319
 CC CC
 QY 898 CTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCGCTGCTGCTGCTGCTG 897
 CC CC
 Db 320 LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerVal 339
 CC CC
 QY 898 GATGAGCTGAC 909
 CC CC
 Db 340 AspGlnAlaAsp 343
 CC CC
 RESULT 9
 AAB58420
 ID AAB58420 standard; protein; 319 AA.
 XX AC AAB58420;
 XX

DT 14-MAR-2001 (first entry)
XX Lung cancer associated polypeptide sequence SEQ ID 758.
DE
XX Human; lung cancer associated protein; neuroprotective; cytoskeletal;
KW cardioactive; immunomodulatory; muscular active; vulnery;
KW gastrocietinal; nephrotropic; antinefctive; gynecological;
KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW proliferative disorder; wound healing; infectious disease.
XX
OS Homo sapiens.
XX PN WO200055180-A2.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000MO-US005918.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
XX PA
XX PI Ruben SM;
XX DR MPI: 2000-587514/55.
XX N-PSDB; AAP18296.
XX PT Lung cancer associated gene sequences, referred to as lung cancer
XX PT such as lung cancer.
XX
XX Claim 11; Page 1275-1276; 1425pp; English.
XX
CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytoskeletal; cardioactive;
CC immunomodulatory; muscular active general; vulnery; gastrocietinal
CC general; nephrotropic; antinefctive; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrocietinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
SQ Sequence 319 AA:

Alignment Scores:
Pred. No.: 2, 89e-167 Length: 319
Score: 194.00 Matches: 194
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.03% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-1 (1-909) x AAB58420 (1-319)
QY 238 GAGAGCAACCTGGCGCTGCTGGGCAACTCTGGCGCTGCTGGCCGCAAGACTGCTG 297
DB 120 GUSerAenLeuArygLeuAlyGlnLeuLeuAryGAlLeuAlyAryGHSAspleuLeu 139
QY 298 CGGCACTGGCGCGGAGCGGCGCGCGCGGCTGCTGCTGCAAGACTTACTTATGACAC 357
DB 140 PchileuAlyAryGlyAryGlyAryGlyAryGlyAryGlyAryGlyAryGlyAryGly 159
QY 358 TCCAGCTCTTCAAG 417

DB 160 SerSerSerSerLyArgThrGlnGlySerCyAArgArgArgGlnSerSerSer 179
QY 418 GCNAATTCTCAGCAGGCTTCAGTGGAGACAGGCTCCCCCAACCAAGCGAGGCGCG 477
DB 180 AIAaenSerGlnGlnGlyGlnTrpGlnTrpGlnTrpGlnTrpGlnTrpGlnTrpGln 199
QY 478 AGTCGGGCGCGCGCGAGT 537
DB 200 SerArgGlyArgProSerGlyGlyAlaArgAArgAArgGlyAlaProAlaAlaPro 219
QY 538 CAGCAGCACTGACAGCGCCGCAAGACTTCTGTGAAGCAAGTGAATCTGTGACATCCG 597
DB 220 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlnGlyValThrCyAspIleArg 239
QY 598 CTCGGGTTTCAGCAGACGATCTGCGACGATGGGCCACTTGTGAGAGGCGCGTGCATCC 657
DB 240 LeuArgValAlaArgAlaGlyTrpCySGlnHISGlyProAlaLeuGlnGlnGlnValAlaSer 259
QY 658 CGAGCGCCCGCAGCGCGCTGGCGGCGAGCTGACGTTGTGGCAGGCGCACCCAGTGTG 717
DB 260 ArgArgProGlnAlaLeuAlaArgGlnLeuAlaPheGlyGlnAlaThrAlaValLeu 279
QY 718 CGCTCAAGGAGACTTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 777
DB 280 ArgSerArgAspLeuGlySerValValCysAspIleLeuPheSerGlnLeuSerTrpLeu 299
QY 778 GAGCGCTTCTGGGCGGCACTTACGAGTGGCGCGCTGCTGTGAC 819
DB 300 AspAlaPheTrpGlyAspTrpLeuSerGlyAlaLeuLeuGln 313

RESULT 10
AAM25705
ID AAM25705 standard; protein; 242 AA.
XX
AC AAM25705;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:1220.
XX
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antitumor; osteopathic; eczema;
KW dermatological; antiallergic; antiaesthetic; antidiabetic; cytoskeletal;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
XX Homo sapiens.
XX PN WO200153455-A2.
XX 26-JUL-2001.
XX
XX 22-DEC-2000; 2000MO-US035017.
XX
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX
XX (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Drmanac RT;
XX PI

Best Local Similarity: 99.25% Mismatches: 0
 Query Match: 39.60% Indels: 1
 DB: 5 Gaps: 0

US-10-030-271-1 (1-909) x ABB06038 (1-361)

```

QY 474 GCGAGTCCGGGCGGCGCCAGTGTGGTGCAGACGGCGGCGAGAGGGCCCGCAGCCGC 533
    |||||
DB 186 AlaGluSerGlyProAlaGlnTrpTrpCysGlnThrAlaIleGluArg-AlaProAlaAl 205

QY 534 ACCCCAGCAGCAGTCAAGACCCCGCAGACCTTCTCTGAAAGGCAAGTCACTGTGACAT 593
    |||||
DB 205 aPpGInGInGInSerGInProAlaArgProSerSerGInGInLysValThrCysAspI 225

QY 594 CCGGCTCCGGGTTGAGAGCAGTACGCGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 653
    |||||
DB 225 eArgLeuAlaValArgAlaGluTrpCysGlnIleGlyProAlaLeuGlnIleGlnValAl 245

QY 654 ATCCCGGCGGCGCCAGGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 713
    |||||
DB 245 aSerArgArgProGlnAlaLeuAlaArgGlnLeuAlaPheGlyGlnAlaThrAlaVa 265

QY 714 GCTGCGCTCAAGGAGCCTGGGCTGTGTGTGTGTGATCAAGTTCTCAGAGCTCTCTCA 773
    |||||
DB 265 IleuArgSerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTy 285

QY 774 TCTGAGCGCTTGTGGGGGCGACTACCTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGG 833
    |||||
DB 285 rLeuAlaPheTrpGlyAspTyLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyVa 305

QY 834 GTTCTGACTGAGGCGCCCTGGCAGAGGCTGTGGCGGCGGCGGAG 873
    |||||
DB 305 lPheLeuThrGlnAlaLeuArgGlnAlaValGlyArgGln 318

RESULT 12
ABB06039
ID ABB06039 standard; protein; 368 AA.
AC ABB06039;
XX
XX 10-MAY-2002 (first entry)
XX
XX Human NS protein sequence SEQ ID NO:131.
XX
XX Human: cytostatic; osteopathic; gynaecological; neuroprotective;
XX antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
XX vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
XX anorectic; muscular; antinfertility; cardiovascular; anticoagulant;
XX anticonvulsant; hypotension; antiasthmatic; immunomodulator; cardiac;
XX gastrointestinal; virocidic; tranquiliser; antidepressant; neuroleptic;
XX contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
XX endometriosis; degenerative disease; multiple sclerosis; psoriasis;
XX rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
XX inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
XX infertility; cardiovascular disease; coagulation disease; hypertension;
XX ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
XX diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
XX gastric ulcer; Alzheimer's disease.
XX
XX Homo sapiens.
XX
XX MO200206315-A2.
XX
XX 24-JUN-2002.
XX
XX 17-JUL-2001; 2001MO-IL000653.
XX
XX 18-JUL-2000; 2000IL-00137345.
XX
XX 15-DEC-2000; 2000IL-00140354.
XX
XX (COMP-) COMPUGEN LTD.
XX

```

PI Mintz J, Freilich S, Bernstein J;
 XX WPI: 2002-155037/20.
 DR N-PSDB; ABL39693.
 DR
 XX
 XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 PT
 PS Claim 6; Page 151-152; 290pp; English.

CC ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virocidic,
 CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antinfertility, cardiovascular,
 CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiac,
 CC immunomodulator, anticonvulsant, antidabetic, tranquiliser, antiulcer,
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
 CC neurotropic and contraceptive activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive

XX
 XX Sequence 368 AA;
 XX

Alignment Scores:
 Pred. No.: 5,68e-100 Length: 368
 Score: 120.00 Matches: 133
 Percent Similarity: 99.25% Conservative: 0
 Best Local Similarity: 99.25% Mismatches: 0
 Query Match: 39.60% Indels: 1
 DB: 5 Gaps: 0

US-10-030-271-1 (1-909) x ABB06039 (1-368)

```

QY 474 GCGAGTCCGGGCGGCGCCAGTGTGGTGCAGACGGCGGCGAGAGGGCCCGCAGCCGC 533
    |||||
DB 193 AlaGluSerGlyProAlaGlnTrpTrpCysGlnThrAlaIleGluArg-AlaProAlaAl 212

QY 534 ACCCCAGCAGCAGTCAAGACCCCGCAGACCTTCTCTGAAAGGCAAGTCACTGTGACAT 593
    |||||
DB 212 aPpGInGInGInSerGInProAlaArgProSerSerGInGInLysValThrCysAspI 232

QY 594 CCGGCTCCGGGTTGAGAGCAGTACGCGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 653
    |||||
DB 232 eArgLeuAlaValArgAlaGluTrpCysGlnIleGlyProAlaLeuGlnIleGlnValAl 252

QY 654 ATCCCGGCGGCGCCAGGCGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 713
    |||||
DB 252 aSerArgArgProGlnAlaLeuAlaArgGlnLeuAlaPheGlyGlnAlaThrAlaVa 272

QY 714 GCTGCGCTCAAGGAGCCTGGGCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTCTCTCA 773
    |||||
DB 272 lLeuArgSerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTy 292

QY 774 TCTGAGCGCTTGTGGGGGCGACTACCTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGG 833
    |||||
DB 292 rLeuAlaPheTrpGlyAspTyLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyVa 312

QY 834 GTTCTGACTGAGGCGCCCTGGCAGAGGCTGTGGCGGCGGCGGAG 873
    |||||
DB 312 lPheLeuThrGlnAlaLeuArgGlnAlaValGlyArgGln 325

RESULT 13
AAE24855

```

ID	AAE24855 standard; protein; 101 AA.
AC	AAE24855;
XX	
DT	22-OCT-2002 (first entry)
XX	
DE	Human DED4 DED (death effector domain) protein.
XX	
KW	Human; death domain; DD; death effector domain; DED; Chlamydia infection; NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis; inflammation; allergy; autoimmunity; allograft rejection; cell division; immune-based pathology; fibrosis; arthritis; graft versus host disease; immunosuppressive; gene therapy; antisense therapy.
XX	
OS	Homo sapiens.
XX	
PN	MO200240680-A2.
XX	
PD	23-MAY-2002.
XX	
PF	15-NOV-2001; 2001MO-US044844.
XX	
PR	17-JUN-2000; 2000US-00715893.
XX	
PR	22-NOV-2001; 2001US-0301889P.
XX	
PA	(BURN-) BURNHAM INST.
XX	
P1	Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
XX	
P1	Stenner-Jlewen F;
XX	
DR	WPI: 2002-500222/53.
XX	
DR	N-PSDB; AABD40075.
XX	
PT	New polypeptide comprising a death domain or death effector domain,
XX	
PT	useful for discovery of drugs that suppress infection, inflammation,
XX	
PT	allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX	
XX	Claim 1; Page 175; 209pp; English.
XX	
PS	The invention relates to an isolated polypeptide comprising a death
XX	
CC	domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX	
CC	is useful for identifying a binding agent, preferably a protein or a drug
XX	
CC	that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX	
CC	domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX	
CC	NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX	
CC	detecting the association of the domain and the candidate binding agent,
XX	
CC	by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX	
CC	chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX	
CC	spectroscopy (MS) and FPA. The invention is useful for modulating the
XX	
CC	level of a cell processes such as cell proliferation, cell adhesion, cell
XX	
CC	stress responses, responses to microbial infection and B cell
XX	
CC	immunoglobulin class switching, in particular apoptosis within a cell.
XX	
CC	Antibody specifically reactive with CTDD DD of C. trachomatis, C.
XX	
CC	muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
XX	
CC	CTDD DD protein is useful for detecting a Chlamydia infection. The
XX	
CC	invention is useful for modulating the activity of oncogenic proteins,
XX	
CC	for treating a pathology caused by the oncogenic proteins and for
XX	
CC	treating bacterial infections by modulating the activity of bacterial
XX	
CC	proteins. The protein and antibody specific for it are useful for
XX	
CC	discovery of drugs that suppress infection, inflammation, allergy,
XX	
CC	sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX	
CC	is useful for treating immune-based pathologies, pathologies associated
XX	
CC	with cell division, inflammatory diseases such as sepsis, fibrosis,
XX	
CC	arthritis, graft versus host disease. The invention is used in antisense
XX	
CC	therapy and gene therapy. The present sequence is human DED4 DED protein
XX	
XX	
SO	Sequence 101 AA;

Alignment Scores:

Pred. No.:	1,1e-51	Length:	101
Score:	67.00	Matches:	67
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match:	22.11%	Indels:	0
DB:	5	Gaps:	0
US-10-030-271-1 (1-909) x AAE24855 (1-101)			
QY	34 TGGAGGAGGATGATGAGCTCTGGACCTAGACCGGATGCTGTGCTTCACCGTATGTCGAG 93		
Db	1 TGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 20		
QY	94 GTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 153		
Db	21 ValValGlyGlyGlyLeuThrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 40		
QY	154 GCTTCCTGGGCGCGCGCGGAGCTTACCCGCGCGCGCGGAGCTTACCGCTTACCGCTTACCGAG 213		
Db	41 AAlpRGGlyAAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGlyLeuLeuGly 60		
QY	214 CTGAGGCGCGCGCGGAGCTGAG 234		
Db	61 LeuGlyArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 67		
RESULT 14			
ID	AAE38898 standard; protein; 101 AA.		
AC	AAE38898;		
XX	18-DEC-2003 (first entry)		
DT			
XX	Human DED4 DED protein.		
DE			
KM	Human; death Domain; DD; death effector domain; DED; cell proliferation;		
KW	Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;		
KW	neural growth factor receptor-interacting death domain; cell adhesion;		
KW	vasotropic; microbial infection; inflammation; allograft rejection; CTDD;		
KW	cell stress response; benign prostatic hypertrophy; antibacterial; NTD;		
KW	apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;		
KW	neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;		
KW	keloid.		
XX			
OS	Homo sapiens.		
XX			
PN	US2003049702-A1.		
PD	13-MAR-2003.		
XX			
PF	15-NOV-2001; 2001US-00001254.		
XX			
PR	17-NOV-2000; 2000US-00715893.		
PR	17-NOV-2000; 2000US-0367360P.		
PR	29-JUN-2001; 2001US-0301889P.		
XX			
PA	(REED/) REED J C.		
PA	(GODZ/) GODZIK A.		
PA	(PAWL/) PAWLOWSKI K.		
PA	(FIOR/) FIORENTINO L.		
PA	(LEES/) LEE S H.		
PA	(ROTH/) ROTH W.		
PA	(STEN/) STENNER-LIEWEN F.		
XX			
PI	Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;		
PI	Stenner-Liewen F;		
XX			
DR	WPI: 2002-500222/53.		
DR	N-PSDB: AAD59057.		
XX			
XX	New polypeptide comprising a death domain or death effector domain,		
PT	useful for discovery of drugs that suppress infection, inflammation,		
PT	allergy, sepsis, autoimmunity, allograft rejection and other diseases.		
XX			
XX	Claim 1; Page 35; 99pp; English.		
CC	The present invention provides novel death Domain (DD) and death effector		

CC domain (DED) proteins and nucleic acids encoding them. The invention also
CC provides death domain containing protein (CTDD) DD and neural growth factor
CC death domain containing protein (CTDD) DD and neural growth factor
CC receptor-interacting death domain (NIDD) DD. The invention is useful for
CC identifying a binding agent (e.g. protein or drug) that binds a DD, DED
CC or NB-ARC domain from DAP3, IRAK4, CTDD, DED4 or NIDD with a candidate
CC binding agent and identifying an effective agent (e.g. protein or drug)
CC that modulates the association of a DD, DED or NB-ARC domain with protein
CC that binds the DD, DED or NB-ARC domain. The invention is also useful for
CC modulating the level of cell process such as apoptosis, cell adhesion,
CC cell proliferation, cell stress responses, responses to microbial
CC infection and B cell immunoglobulin class switching. Dds, Deds and NB-ARC
CC domains and/or anti-DD, anti-DDD or anti-NB-ARC domain antibodies are
CC useful for discovery of drugs that suppress infection, autoimmunity,
CC inflammation, allergy, allograft rejection, sepsis and other diseases.
CC DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
CC autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
CC hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
CC inflammatory hyperplasia and smooth muscle cell proliferation in arteries
CC following balloon angioplasty (restenosis). The invention is also used in
CC antibody therapy and gene therapy. The present sequence is human DED4 DED
CC protein. The DED4 gene is located on chromosome 1

XX Sequence 101 AA;

SQ Alignment Scores:

Pred. No.:	1,1e-51	Length:	101
Score:	67.00	Matches:	67
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	22.11%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-1 (1-909) x AAM38898 (1-101)

QY 34 TGGAGAGAGATGAGTCTGAGCTACTACGAGATGCTGCTTACCGTATGTTGAG 93

DB 1 TTPGLUGLUAAPGLUCYSLEUAAPRYTYRGLYMECTUSENSEUHSISARGMETPHEGLU 20

QY 94 GTGGTGGGGGGGAGCAATGACCCGAGCTGAGCTGCTGCTTCTTCTGCTGAGTGA 153

DB 21 ValValGlyGlyGlnLeuThrGlnCysGlnLeuGlnLeuAlaPheLeuLeuAaspGln 40

QY 154 GCTCTGAGCGCCGCGGAGGCTTAGCCCGGAGCCGCGAGCGGCTAGAGCTCTGTGAG 213

DB 41 AAlaprogLYAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGlnLeuLeuGln 60

QY 214 CTGGAGCGCCGCGGAGTGC 234

DB 61 LeuGluArgArgGlyGlnCys 67

RESULT 15

AAM41591

ID AAM41591 standard; protein; 217 AA.

XX AAM41591;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6522.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX peripheral nervous system; neuropathy; central nervous system; CNS;

XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX chemokine; thrombolytic; drug screening; arthritis; inflammation;

XX Leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JUN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00553317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB; AAI60747.

XX Example 2; SEQ ID NO 6522; 10078bp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the

XX encoded polypeptides (AAM38842-AAM42213) with nootropic,

XX immunosuppressant and cytostatic activity. The polynucleotides are useful

XX in gene therapy. A composition containing a polypeptide or polynucleotide

XX of the invention may be used to treat diseases of the peripheral nervous

XX system, such as peripheral nervous injuries, peripheral neuropathy and

XX localized neuropathies and central nervous system diseases, such as

XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX utilisation of the activities such as: Immune system suppression,

XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic

XX and thrombolytic activity, cancer diagnosis and therapy, drug screening, and

XX assays for receptor activity, arthritis and inflammation, leukaemia and

XX C.N.S disorders. Note: The sequence data for this patent did not form

XX part of the printed specification

XX SQ Sequence 217 AA;

XX Alignment Scores:

Pred. No.:	7.95e-51	Length:	217
Score:	66.00	Matches:	66
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	21.78%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-1 (1-909) x AAM41591 (1-217)

QY 238 GAGAGCACTGCGGCTGAGGCACTCTGCGGCTGCTGCGCCGAGCACTGCTG 297

DB 109 GluSerAsnLeuArgLeuLeuGlnLeuLeuAlaValLeuAlaArgHisAspLeuLeu 128

QY 298 CCGACCTGGCGCGGCAAGCGCGCGCGCAAGTCTCTCCAGAACGCTATAGTGCACC 357

DB 129 ProHisLeuAlaArgLysArgArgProValSerProGluArgTyrSerTyrGlyThr 148

QY 358 TCGAGCTCTTCAAGAGACAGAGGGTACTGCGCGCGCTGCGAGTCAAGCACTTCT 417

DB 149 SerSerSerSerLysArgThrGlnGlySerCysArgArgArgArgGlnSerSerSer 168

QY 418 GCAATTCTCAGCAGGGT 435

DB 169 AlaAsnSerGlnGlnGly 174

RESULT 16

AAM39805

CC Inhibitors of the polypeptides can be used as apoptotic agents
XX Sequence 318 AA;

Alignment Scores:

Pred. No.:	0.012	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-1 (1-909) x AAW90109 (1-318)

OY 583 ACCTGACATCCGGCTCGGTTGAGACAGACTACTGC 621
DB 192 TnCyAspIleArgLeuArgValArgAlaGluTyrCys 204

RESULT 18

AAW90109
ID AAW90109 standard; protein; 318 AA.

AC AAW90109;

DT 12-APR-1999 (first entry)

DE Mouse FLAME-2 protein.

XX FLAME-2; FADD-like apoptotic/anti-apoptotic molecule; mouse; apoptosis;
KW HIV; infection; Alzheimer's disease; cancer; therapy.

OS Mus sp.

PN W09852963-A1.

PD 26-NOV-1998.

PF 20-MAY-1998; 98WO-US010200.

PR 20-MAY-1997; 97US-00859167.

PA (UYJB-) UNIV JEFFERSON THOMAS.

PI Alnemr1 ES;

DR WPI; 1999-045296/04.

DR N-PSDB; AAV74138.

XX New isolated FADD-like anti-apoptotic molecules - used to develop
PT apoptotic and anti-apoptotic agents for treating, e.g. HIV infection,
PT Alzheimer's disease or neoplastic conditions.

XX Example; Page 44-45; 68pp; English.

CC This is the amino acid sequence of mouse FLAME-2, or FADD-like
CC apoptotic/anti-apoptotic molecule 2. FLAME-2 is a novel anti-apoptotic
CC protein that is similar in structure to FADD, but its C-terminal region
CC does not have death domain homology. Its structure suggests that it could
CC be an adaptor molecule for an as yet unidentified signalling complex.
CC FLAME-2 interacts weakly with Mch4 and Mch5 but does not interact with
CC FADD. It can abrogate UV-induced apoptosis and to a lesser degree inhibit
CC Fas/TNFR-induced apoptosis in the same cell line. The FLAME-2 amino acid
CC sequence was deduced from the nucleotide sequence of an isolated EST (see
CC AAV74138). Human FLAME-1 and FLAME-2 proteins (see AAW90107-08) are
CC claimed. Host cells, recombinant vectors, and methods of using FLAME to
CC identify substrates, activators or inhibitors of FLAME are provided.
CC FLAME-1, FLAME-2 and agonists can be used to inhibit apoptosis, e.g. for
CC treating HIV infection or Alzheimer's disease. Inhibitors of the
CC polypeptides can be used as apoptotic agents

XX Sequence 318 AA;

Alignment Scores:

Pred. No.:	0.012	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-1 (1-909) x AAW90109 (1-318)

OY 583 ACCTGACATCCGGCTCGGTTGAGACAGACTACTGC 621
DB 192 TnCyAspIleArgLeuArgValArgAlaGluTyrCys 204

RESULT 19

AAV51023
ID AAV51023 standard; protein; 318 AA.

AC AAV51023;

DT 17-MAR-2000 (first entry)

DE Murine DEDD protein.

XX DEDD; murine; apoptosis; protein biosynthesis inhibitor; CD95.

OS Mus sp.

PN DE19825621-A1.

PD 09-DEC-1999.

PF 08-JUN-1998; 98DE-01025621.

PR 08-JUN-1998; 98DE-01025621.

PA (DEKR-) DEUT KREBSFORSCHUNGSGEZENTRUM.

PI Peter M. Kramer P;

DR WPI; 2000-063547/06.

DR N-PSDB; AA243924.

XX New protein for regulating apoptosis, particularly for diagnosis or
PT treatment of e.g. tumors and acquired immune deficiency syndrome.

XX Claim 7; Fig 1B; 12pp; German.

CC This invention describes a novel DEDD protein (I) for regulating
CC apoptosis which has antitumor and antiviral activity. (I) is expressed
CC ubiquitously and after induction of the CD95 apoptotic signalling pathway
CC migrates to the nucleus/nucleoli where it inhibits transcription of
CC ribosomal DNA, and thus biosynthesis of proteins, including those with
CC anti-apoptotic activity. (I), or the DNA (II), encoding it, is used for
CC regulation and diagnostic detection of apoptosis, particularly in cases
CC of disease, e.g. of the immune system (e.g. acquired immune deficiency
CC syndrome) and tumors, also for studying apoptosis or its regulation.
CC Antibodies (Ab) specific for (I) are used to detect/quantify (I). (I) is
CC particularly as a function of time, also for inhibition of (I). (I) is
CC used to raise Ab and for detection of specific autoantibodies. (II), or
CC derived primers, can be used to detect expression and organization of the
CC corresponding gene, also for inducing expression of (I) in vivo or in
CC selected tissues. This sequence represents the murine DEDD protein
CC described in the method of the invention

XX Sequence 318 AA;

Alignment Scores:

Pred. No.:	0.012	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	4.29%	Indels:	0
DB:	3	Gaps:	0

US-10-030-271-1 (1-909) x AAYS1023 (1-318)

QY 583 ACCTGTGACATCCGGCTCCGGTTCCAGACAGACTACTGC 621
 |||
 DB 192 ThrcYAspIleArgLeuArgValArgAlaGluTyr-Cys 204

RESULT 20

AAYS1022
 ID AAYS1022 standard; protein; 318 AA.

AC AAYS1022;

DT 17-MAR-2000 (first entry)

DE Human DEDD protein.

KW DEDD; human; apoptosis; protein biosynthesis inhibitor; CD95.

OS Homo sapiens.

PN DE19825621-A1.

PD 09-DEC-1999.

PF 08-JUN-1998; 98DE-01025621.

PR 08-JUN-1998; 98DE-01025621.

PI (DEKR-) DEUT KREBSFORSCHUNGSENTRUM.

PI Peter M, Kramer P;

XX WPI; 2000-063547/06.

DR N-PSDB; AA243923.

PT New protein for regulating apoptosis, particularly for diagnosis or

PS treatment of e.g. tumors and acquired immune deficiency syndrome.

PS Claim 1; Fig 1A; 12pp; German.

CC This invention describes a novel DEDD protein (I) for regulating
 CC apoptosis which has antitumor and antiviral activity. (I) is expressed
 CC ubiquitously and after induction of the CD95 apoptotic signalling pathway
 CC migrates to the nucleus/nucleolus where it inhibits transcription of
 CC ribosomal DNA, and thus biosynthesis of proteins, including those with
 CC anti-apoptotic activity. (II), or the DNA (II), encoding it, is used for
 CC regulation and diagnostic detection of apoptosis, particularly in cases
 CC of disease, e.g. of the immune system (e.g. acquired immune deficiency
 CC syndrome) and tumors; also for studying apoptosis or its regulation.
 CC Antibodies (Ab) specific for (I) are used to detect/quantify (I).
 CC particularly as a function of time, also for inhibition of (I). (I) is
 CC used to raise Ab and for detection of specific autoantibodies. (II), or
 CC derived primers, can be used to detect expression and organization of the
 CC corresponding gene, also for inducing expression of (I) in vivo or in
 CC selected tissues. This sequence represents the human DEDD protein
 CC described in the method of the invention

XX SQ Sequence 318 AA;

Alignment Scores:

Pred. No.: 0.012 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.29% Indels: 0
 DB: 3 Gaps: 0

US-10-030-271-1 (1-909) x AAYS1022 (1-318)

QY 583 ACCTGTGACATCCGGCTCCGGTTCCAGACAGACTACTGC 621
 |||
 DB 192 ThrcYAspIleArgLeuArgValArgAlaGluTyr-Cys 204

RESULT 21

AAB93016
 ID AAB93016 standard; protein; 318 AA.

AC AAB93016;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11773.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

PI (HELI-) HELIX RES INST.

PI Ota T, Isega T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PS length cDNAs defined in the specification, and for the detection and/or

PT diagnosis of the abnormality of the proteins encoded by the full-length

PS cDNAs.

PS Claim 8; SEQ ID NO 11773; 2537bp + Sequence Listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 318 AA;

Alignment Scores:
 Pred. No.: 0.012 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.29% Indels: 0
 DB: 4 Gaps: 0

US-10-030-271-1 (1-909) x AAB93016 (1-318)

QY 583 ACCTGTGACATCCGGCTCCGGGTTGAGACAGACTACTGC 621
 Db 192 TnTCyASpILeArGLenArGLValArgAlaGluTyrCys 204

RESULT 22
 AAB94040

ID AAB94040 standard; protein; 318 AA.

XX AAB94040;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:14195.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-00116126.

PR 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99UP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INSTR.

XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PS Claim 8; SEQ ID NO 14195; 2537bp + Sequence listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-

CC length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the

CC complementary strand of a polynucleotide which comprises one of the 5602

CC nucleotide sequences defined in the specification; where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence; where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in

CC gene therapy. The primers are useful for synthesizing polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893

CC represent human amino acid sequences; and AAH13629 to AAH1632 represent

CC oligonucleotides, all of which are used in the exemplification of the

CC present invention

XX Sequence 318 AA;

XX SQ

Alignment Scores:

Pred. No.: 0.012 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.29% Indels: 0
 DB: 4 Gaps: 0

US-10-030-271-1 (1-909) x AAB94040 (1-318)

QY 583 ACCTGTGACATCCGGCTCCGGGTTGAGACAGACTACTGC 621
 Db 192 TnTCyASpILeArGLenArGLValArgAlaGluTyrCys 204

RESULT 23
 AAE26087

ID AAE26087 standard; protein; 318 AA.

XX AAE26087;

DT 14-NOV-2002 (first entry)

DE Human FLAME-2 protein.

XX Human; FADD-like apoptotic/anti-apoptotic protein; Alzheimer's disease;

XX gene therapy; human immunodeficiency virus; HIV infection; apoptosis;

XX FLAME-2.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Region 1..22 /note="IE-2 region"

XX Region 23..101 /note="FADD-DED-Homology region; N-terminal region"

XX Domain 102..318 /note="C-terminal domain"

XX US2002086963-A1.

XX 04-UTL-2002.

XX 22-AUG-2001; 2001US-00935223.

XX 28-OCT-1997; 97US-00959167.

XX 26-MAR-1999; 99US-00276993.

XX 28-NOV-2000; 2000US-00723450.

XX (UYDE-) UNITV JEFFERSON THOMAS.

XX Alnemri ES;

XX WPI; 2002-642259/69.

XX N-PSDB; AAD43203.

XX Novel FADD-like apoptotic/anti-apoptotic proteins useful for inhibiting

XX apoptosis, treating diseases characterized by apoptosis e.g. HIV

XX infection and Alzheimer's disease, and for identifying modulators of the

XX protein.

XX Claim 1; Page 16-17; 35pp; English.

XX The invention relates to FADD-like apoptotic/anti-apoptotic proteins

XX (FLAME 1 or 2) and nucleic acid molecules encoding such proteins. FLAME

XX sequences are useful for inhibiting apoptosis and for gene therapy of

XX diseases characterised by apoptosis including HIV infection and

XX Alzheimer's disease. FLAME inhibitors are useful as apoptotic agents and

XX activators are useful as anti-apoptotic agents. FLAME-1 is useful as a

XX substrate for caspase in assays to identify caspase inhibitors. The

XX present sequence is human FLAME-2 protein

XX Sequence 318 AA;

XX SQ

Alignment Scores:

Pred. No.: 0.012 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.29% Indels: 0
 DB: 5 Gaps: 0

US-10-030-271-1 (1-909) x AAE26087 (1-318)

QY 583 ACCTGTGACATCCGGCTCCGGGTTGAGCAGAGTACTGC 621
 DB 192 ThrcysAspIleArgLeuArgValArgAlaGluTyrCys 204

RESULT 24

AAE26088
 ID AAE26088 standard; protein; 318 AA.

AC AAE26088;

DT 14-NOV-2002 (first entry)

DE Mouse FLAME-2 protein.

KM Mouse; FADD-like apoptotic/anti-apoptotic protein; Alzheimer's disease;
 gene therapy; human immunodeficiency virus; HIV infection; apoptosis;

KW FLAME-2.

XX Mus sp.

PN US2002086983-A1.

PD 04-JUL-2002.

PF 22-AUG-2001; 2001US-00935223.

PR 28-OCT-1997; 97US-00959167.

PR 26-MAR-1999; 99US-00276993.

PR 28-NOV-2000; 2000US-00723450.

PA (UYUE-) UNITV JEFFERSON THOMAS.

PI Alnemri ES;

XX MPI: 2002-642259/69.

DR N-PSDB; AAD43204.

XX Novel FADD-like apoptotic/anti-apoptotic proteins useful for inhibiting

PT apoptosis, treating diseases characterized by apoptosis e.g. HIV

PT infection and Alzheimer's disease, and for identifying modulators of the

PT protein.

XX Disclosure; Page 19; 35pp; English.

PS The invention relates to FADD-like apoptotic/anti-apoptotic proteins

XX (FLAME 1 or 2) and nucleic acid molecules encoding such proteins. FLAME

CC sequences are useful for inhibiting apoptosis and for gene therapy of

CC diseases characterized by apoptosis including HIV infection and

CC Alzheimer's disease. FLAME inhibitors are useful as apoptotic agents and

CC activators are useful as anti-apoptotic agents. FLAME-1 is useful as a

CC substrate for caspase in assays to identify caspase inhibitors. The

CC present sequence is mouse FLAME-2 protein

XX SQ Sequence 318 AA;

Alignment Scores:

Pred. No.: 0.012 Length: 318

Score: 13.00 Matches: 13

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 4.29% Indels: 0

DB: 5 Gaps: 0

US-10-030-271-1 (1-909) x AAE26088 (1-318)

QY 583 ACCTGTGACATCCGGCTCCGGGTTGAGCAGAGTACTGC 621
 DB 192 ThrcysAspIleArgLeuArgValArgAlaGluTyrCys 204

RESULT 25

AAU41693
 ID AAU41693 standard; protein; 104 AA.

AC AAU41693;

DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #2589.

KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR MPI: 2001-616774/71.

XX N-PSDB; AAS59515.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

PT vaccinating against and diagnosing infections, especially useful for

PT treating acne vulgaris.

XX Example 1; SEQ ID NO 2888; 1069pp; English.

XX Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic

CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by

CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central

CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the

CC presence or absence of P. acnes in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention

CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and

CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by

CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for

CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 104 AA;

Alignment Scores:

Pred. No.: 7.57 Length: 104

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0

CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 51 AA;

SO Alignment Scores:

Pred. No.:	68.3	Length:	51
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.00%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-1 (1-909) x AAU41040 (1-51)

OY 146 AGCAGAAAGCCGAGAGCTCCAGCTCG 120

DB 40 SerArgLysAlaArgSerSerSer 48

RESULT 30

ABM37559

ID ABM37559 standard; protein; 51 AA.

AC ABM37559;

DT 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #2235.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO2003033515-A1.

PD 24-APR-2003.

PF 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

PA (CORI-) CORIXA CORP.

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglas J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64442.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 2235; 1481bp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

CC encoding a Propionibacterium acnes protein. The invention also relates to

CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to

CC immunogenic fragments of P. acnes polypeptides. The invention

CC additionally encompasses expression vectors and host cells comprising a

CC polynucleotide of the invention; antibodies against polypeptides of the

CC invention; fusion proteins comprising a polypeptide of the invention; a

CC method for stimulating an immune response specific for a P. acnes

CC polypeptide and an isolated T cell population comprising T cells prepared

CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 51 AA;

SO Alignment Scores:

Pred. No.:	68.3	Length:	51
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.00%	Indels:	0
DB:	6	Gaps:	0

US-10-030-271-1 (1-909) x ABM37559 (1-51)

OY 146 AGCAGAAAGCCGAGAGCTCCAGCTCG 120

DB 40 SerArgLysAlaArgSerSerSer 48

RESULT 31

ABM51895

ID ABM51895 standard; protein; 51 AA.

AC ABM51895;

DT 20-OCT-2003 (first entry)

DE Propionibacterium acnes predicted ORF-encoded polypeptide #16571.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO2003033515-A1.

PD 24-APR-2003.

PF 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

PA (CORI-) CORIXA CORP.

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes WJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglas J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64498.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1; SEQ ID NO 16571; 1481bp; English.

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
CC XX

Sequence 51 AA:
SQ

Alignment Scores:
Pred. No.: 68.3 Length: 51
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0

US-10-030-271-1 (1-909) x ABM51895 (1-51)

OY 227 CCGCGGCGCTCCAGCTCCAGCAGAGC 201
DB ||||| 25 ProArgArgSerSerSerArgSer 33

RESULT 32
ADN01127
ID ADN01127 standard; protein; 68 AA.
XX
AC ADN01127;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human cell growth, differentiation, and death-associated protein #11.
XX
KW human, cell growth; cell differentiation; cell death; CGGD;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis; bursitis;
XX cirrhosis; hepatitis; polycythaemia vera; psoriasis; cancer;
XX developmental disorder; Cushing's syndrome; hypothyroidism;
XX neurological disorder; epilepsy; stroke; Alzheimer's disease;
XX Pick's disease; Huntington's disease; Parkinson's disease;
XX multiple sclerosis; autoimmune disorder; inflammatory disorder; AIDS;
KW allergy; anaemia; asthma; contact dermatitis; diabetes mellitus;
KW reproductive disorder; infertility; endometriosis; uterine fibroid.
XX
OS Homo sapiens.
XX
PN WO2004031364-A2.
XX
PD 15-APR-2004.
XX
PF 03-OCT-2003; 2003WO-US031441.
XX
PR 03-OCT-2002; 2002US-0416205P.

PR 25-OCT-2002; 2002US-0421521P.
PR 21-NOV-2002; 2002US-0428376P.
PR 23-DEC-2002; 2002US-0436258P.
PR 10-JAN-2003; 2003US-0439292P.
PR 13-FEB-2003; 2003US-0447578P.
XX
PA (INCY-) INCYTE CORP.
PA (BURL/) BURLILL J D.
XX
PI Elliott VS, Swarnakar A, Tang YT, Yue H, Lu DM, Khare R;
PI Charva NK, Richardson TW, Margulis JP, Lal PG, Nguyen DB, Lee SY;
PI Tran UK, Bhatia UC, Lee S, Blake JU, Ho A, Zheng W, Gao J, Tran B;
PI Yang YG, Gietzen KU, Kafalia AUA,
XX
DR WPI; 2004-330172/30.
DR N-PSDB; ADN01152.
XX
PT New isolated polypeptides associated with cell growth, differentiation
PT and death, useful for diagnosing, treating or preventing e.g.
PT atherosclerosis, psoriasis, cancers, Alzheimer's disease, AIDS, anemia,
PT diabetes mellitus or infertility.
XX
PS Claim 1, SEQ ID NO 11; 213pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC proteins that are associated with cell growth, differentiation, and death
CC (CGGD). The DNA and protein sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of CGGD, such as: cell proliferative disorders (e.g.
CC arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis,
CC polycythaemia vera, psoriasis and cancer), developmental disorders (e.g.
CC Cushing's syndrome and hypothyroidism), neurological disorders (e.g.
CC epilepsy, stroke, Alzheimer's disease, Pick's disease, Huntington's
CC disease, Parkinson's disease and multiple sclerosis),
CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, anaemia, asthma,
CC contact dermatitis and diabetes mellitus), and reproductive disorders
CC (e.g. infertility, endometriosis and uterine fibroid) The present amino
CC acid sequence represents a human CGGD-associated protein of the
CC invention.
CC XX

Sequence 68 AA:
SQ

Alignment Scores:
Pred. No.: 65.5 Length: 68
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0

US-10-030-271-1 (1-909) x ADN01127 (1-68)

OY 243 GCTCTGCGCGACATGCGCGGCGCTC 217
DB ||||| 18 AlaLeuAlaAlaLeuProAlaAlaLeu 26

RESULT 33
AAU41429
ID AAU41429 standard; protein; 71 AA.
XX
AC AAU41429;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #2325.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX

XX	MO200181581-A2.
XX	01-NOV-2001.
XX	20-APR-2001; 2001WO-US012865.
XX	21-APR-2000; 2000US-0199047P.
PR	02-JUN-2000; 2000US-0208841P.
PR	07-JUL-2000; 2000US-0216747P.
XX	(CORI-) CORIXA CORP.
PA	
XX	Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI	L'maisonmeuve J, Zhang Y, Jen S, Carter D;
PI	
XX	WPI; 2001-616774/71.
DR	N-PSDB; AAS59515.
XX	
PT	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris.
XX	
PS	Example 1; SEQ ID NO 2624; 1069pp; English.
XX	
CC	Sequences AAU9105-AAU6017 represent Propionibacterium acnes immunogenic
CC	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	downregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC	this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
XX	Sequence 71 AA;
XX	
XX	Alignment Scores:
XX	Pred. No.:
XX	Score: 65.1 Length: 71
XX	Percent Similarity: 9.00 Matches: 9
XX	Best Local Similarity: 100.00% Conservativity: 0
XX	Query Match: 100.00% Mismatches: 0
XX	Indels: 0
XX	Gaps: 0
XX	
XX	US-10-030-271-1 (1-909) x AAU41429 (1-71)
OY	36 CCAGCAGCGGCGCCGGGTCGACCCGGA 10
DB	
DB	6 PROAIAATGGLYATGGLYATGGLYATGGLY 14
XX	
XX	RESULT 34
XX	ABM37948
XX	ABM37948 standard; protein; 71 AA.
XX	
XX	ABM37948;
XX	
XX	20-OCT-2003 (first entry)
XX	
DE	Propionibacterium acnes predicted ORF-encoded polypeptide #2624.
XX	
XX	Acne vulgaris; antisborrheic; dermatological; antibacterial;
XX	immunostimulant; immune response; vaccine.
KW	

XX Proionibacterium acnes.
XS WO2003033515-A1.
XN 24-APR-2003.
XP 11-OCT-2002; 2002MO-USO32727.
XX 15-OCT-2001; 2001US-00978825.
PR (CORI-) CORIXA CORP.
PA Micham JL, Skeiky YAM, Persing DH, Bhata A, Maisonneuve JI;
PJ Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vailleve-Douglas J;
XX WPI: 2003-381789/36.
DR N-PsDB; ACF64444.

New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PP or for stimulating an immune response specific for a P. acnes protein.

Example 1; SEQ ID NO 2624; 1481bp; English.

The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WHO at ftp.wipo.int/pub/published_pct_sequences

SO Sequence 71 AA;

Alignment Scores:		
Pred. No.: 65.1	Length:	71
Score: 9.00	Matches:	9
Percent Similarity: 100.00%	Conservative: 0	
Best Local Similarity: 100.00%	Mismatches: 0	
Query Match: 3.00%	Indels: 0	
DB: 6	Gaps: 0	

US-10-030-271-1 (1-909) x ABM37948 (1-71)

```

QY      36 COAGCAGCGGACC CGGGATCGAACCCGGA   10
        |||||              |||||
Db       6 ProAlaArgGIyArGgIAyrgProclly  14
    
```

RESULT 35
AAU62949 standard; protein; 89 AA.
ID

```
XX AC AU62949;
XX XX
XX DT 27-FEB-2002 (first entry)
XX XX
XX DE Propionibacterium acnes immunogenic protein #23845.
XX XX
XX KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX KW dermatological; osteopathic; neuroprotectant.
XX OS
XX XX
XX PN WO200181581-A2.
XX XX
XX PD 01-NOV-2001.
XX XX
XX PF 20-APR-2001; 2001WO-US012865.
XX XX
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR N-PSDB; AAS59630.
XX XX
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX XX
XX PS Example 1; SEQ ID NO 24144; 1069bp; English.
XX XX
XX CC Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 89 AA;
XX XX
XX Alignment Scores:
XX Pred. No.: 62.9 Length: 89
XX Score: 9.00 Matches: 9
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 3.00% Indels: 0
XX DB: 4 Gaps: 0
XX XX
US-10-030-271-1 (1-909) x AU62949 (1-89)
QY 227 CCGCGCGCTCCAGCTCCAGCAGAGC 201
|||||
```

```
DB 66 ProArgArgSerSerSerArgSer 74
RESULT 36
ABMS9468
ID ABMS9468 standard; protein; 89 AA.
XX
XX AC ABMS9468;
XX XX
XX DT 20-OCT-2003 (first entry)
XX XX
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #24144.
XX XX
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine.
XX XX
XX OS
XX XX
XX PN WO2003033515-A1.
XX XX
XX PD 24-APR-2003.
XX XX
XX PF 11-OCT-2002; 2002WO-US032727.
XX XX
XX PR 15-OCT-2001; 2001US-00978825.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX PI Barsh B, Valliave-Douglas J;
XX DR WPI; 2003-381789/36.
XX DR N-PSDB; ACF64559.
XX XX
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX XX
XX PS Example 1; SEQ ID NO 24144; 1481bp; English.
XX XX
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention; antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
XX CC and the kit is useful for performing a diagnostic assay. The present
XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
XX CC reading frame) contained within the P. acnes polynucleotides of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 89 AA;
XX XX
XX Alignment Scores:
XX Pred. No.: 62.9 Length: 89
XX Score: 9.00 Matches: 9
XX XX
```

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.00%
DB: 6
US-10-030-271-1 (1-909) x ABM59468 (1-89)
OY 227 CCGCGGCGCTCCAGCTCCAGCAGACC 201
DB 66 ProArgArgSerSerSerSerArgSer 74
RESULT 37
ABO81797
ID ABO81797 standard; protein; 91 AA.
XX ABO81797;
AC ABO81797;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #13972.
XX Pseudomonas aeruginosa polypeptide #13972.
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX US6551795-B1.
XX 22-APR-2003.
XX 18-FEB-1999; 99US-00252991.
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1996; 98US-0094190P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Rubenfeld MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX DR N-PSDB; ABD15368.
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 30543; 455bp; English.
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biotech technology. Sequences ABO67826-
XX CC ABO81396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
SQ Sequence 91 AA;
Alignment Scores:
Pred. No.: 62.7 Length: 91
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 7 Gaps: 0

US-10-030-271-1 (1-909) x ABO81797 (1-91)
OY 505 AGACGCGCGCGAGAGAGCGCCCGCAGCC 531
DB 5 ArgArgArgArgArgGlyAlaProAla 13
RESULT 38
AAU46695
ID AAU46695 standard; protein; 125 AA.
XX AAU46695;
AC AAU46695;
XX 27-FEB-2002 (first entry)
XX Propionibacterium acnes immunogenic protein #7591.
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertrophic; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX (CORI-) CORIXA CORP.
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
XX DR N-PSDB; AAS59534.
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX Example 1; SEQ ID NO 7890; 1069bp; English.
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hypertrophic and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 125 AA;
Alignment Scores:
Pred. No.: 59.8 Length: 125

PT preventing, diagnosing and/or treating e.g. Gaucher's disease,
PT Alzheimer's disease, Scimitar syndrome, Creutzfeldt-Jacob disease,
PT diabetes mellitus and multiple sclerosis.

PS Claim 11; Page 480-481; 546pp; English.

CC AAD07651-AA0D7645 represent cDNAs corresponding to 25 human secreted
CC protein genes, and AAEO3052-ABE03126 represent the proteins they encode
CC AAE03127-AAE03150 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 25 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angioinogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention

XX
XX Sequence 211 AA;

SQ Sequence 211 AA;

Alignment Scores:		
Pred. No.:	55.4	length: 211
Score:	9.00	Matches: 9
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	3.00%	Indels: 0
DB:	4	Gaps: 0

US-10-030-271-1 (1-909) X AAE03094 (1-211)

QY 688 CCAAGTGGCCGGCCAGCGCCTGGGGCC 662

Db 126 ProIaIaIaIaProIaIaProIaIa 134

RESULT 41

ID ABG63622 standard; protein; 211 AA.

AC ABG63622;

DT 27-AUG-2002 (first entry)

DE Human albumin fusion protein #297.

KM Albumin fusion protein; therapeutic protein X; human albumin; HA
 KM human serum albumin; HSA; cancer; reproductive disorder;
 KM digestive disorder; immune disorder; endocrine disorder;
 KM haematopoietic disorder; neural disorder; connective disorder;
 KM cytotoxic; antineoplastic; antidiabetic; haemostatic; nootropic;
 KM immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
 KM neuroprotective; antibacterial; antimicrobial; neuroleptic;
 KM osteopathic; antiallergic

OS Homo sapiens.
OS Synthetic.

OS Synthetic.

XX WO200177137-A1
PN

PD 18-OCT-2001

PF 12-APR-2001; 2001WO-US011988.

PR 12-APR-2000; 2000US-0229358P

PR 21-DEC-2000; 2000US-0256931P

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Haseltine WA;

DR WPI; 2002-010886/01.

PT New fusion protein for treating disease e.g. diabetes comprises an albumin fused to a therapeutic protein.

PS Claim 1; Page 758; 2102pp; English

CC The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
CC schizophrenia), and connective tissue disorders (e.g. osteoporosis, arthritis).
CC ABG63326-ABG6518 represent albumin fusion proteins of the invention
XX
XX Sequence 211 AA;

SQ Sequence 211 AA;

Alignment Scores:	
Pred. No.:	55.4
Score:	9.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.00%
DB:	5
Gaps:	0
length:	21
Matches:	9
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-030-271-1 (1-909) X ABG63622 (1-211)

QY	688	CCAGCTGCCCGCCAGCGCCTGGGGCC	662
Db	126	ProIaIaIaProIaProIaIa	134

RESULT 42

ID ADAS7327 standard; protein; 211 AA.

AC ADA57327;

DT 20-NOV-2003 (first entry)

DE Human secreted protein #39.

immunosuppressive; antiinflammatory; antisthmatic; antiallergic;
 cytosolic; cerebroprotective; neuroprotective; nootropic;
 cardiovascular; antihypertensive; gene therapy;
 human secreted protein; immune disorder; inflammation;
 respiratory disorder; cancer; CNS disorder; neurodegenerative disorders
 inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy
 multiple sclerosis; ischaemic brain injury; Parkinson's disease;
 Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 triple helix formation; antisense gene therapy; forensic biology.

XX	Homo sapiens
OS	

OS Homo sapiens

XX WO2002102994-A2.
 XX
 XX 27-DEC-2002.
 XX
 XX 19-MAR-2002; 2002WO-US008278.
 XX
 XX 21-MAR-2001; 2001US-0277340P.
 XX 19-JUL-2001; 2001US-0306171P.
 XX 13-NOV-2001; 2001US-0331287P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX WPI; 2003-167512/16.
 XX N-PSDB; ADA56431.
 XX
 XX New human secreted polypeptides and polynucleotides, useful for
 XX diagnosing, treating or preventing e.g. immune disorders, inflammatory
 XX conditions, respiratory disorders, cancers, CNS disorders, or
 XX neurodegenerative disorders.
 XX
 XX Claim 13; SEQ ID NO 1517; 1754pp; English.
 XX
 XX The invention relates to 592 new human secreted polypeptides useful for
 XX diagnosing, treating or preventing e.g. immune disorders, inflammatory
 XX conditions, respiratory disorders, cancers, CNS disorders, or
 XX neurodegenerative disorders, or polypeptides comprising an amino acid
 XX sequence at least 95% identical to the new sequences. The polypeptides,
 XX antibodies or antibody fragments that bind to the polypeptides, nucleic
 XX acids encoding the polypeptides, agonists or antagonists that binds to
 XX the polypeptide, are useful in preparing diagnostic or pharmaceutical
 XX compositions for diagnosing, treating or preventing an e.g. immune
 XX disorders, inflammatory conditions (e.g. inflammatory bowel disease,
 XX nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 XX allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 XX (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
 XX disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 XX cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 XX polynucleotides are useful for chromosome identification, chromosome
 XX mapping, for controlling gene expression through triple helix formation
 XX or antisense DNA or RNA, in gene therapy, for identifying individuals
 XX from minute biological samples, in forensic biology, and as hybridization
 XX probes. The polypeptides are useful for as molecular weight markers on
 XX sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
 XX gels, to raise antibodies, for testing biological activities, and for
 XX treating or preventing neural disorders, immune system disorders,
 XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 XX renal, proliferative and/or cancerous diseases. This sequence corresponds
 XX to one of the polypeptide of the invention. Note: The sequence data for
 XX in this patent did form part of the printed specification, but was obtained
 XX in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX SO Sequence 211 AA;
 XX
 XX Alignment Scores:
 XX Pred. No.: 55.4 Length: 211
 XX Score: 9.00 Matches: 9
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 3.00% Indels: 0
 XX DB: Gaps: 0
 XX
 XX US-10-030-271-1 (1-909) x ADA57327 (1-211)
 QY 688 CCAGCTGCGCGCGCCAGCGCTGCGGCC 662
 Db 126 ProAlaAlaAlaProAlaProGlyAla 134
 RESULT 43
 ADL76887

ID ADL76887 standard; protein; 211 AA.
 XX
 XX AC ADL76887;
 XX
 XX 20-MAY-2004 (first entry)
 XX
 XX Albumin fusion protein related therapeutic protein X, SEQ ID NO 369.
 XX
 XX albumin fusion protein; cytosolic; antianaemic; antiarthritic;
 XX antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
 XX antipsoriatic; antibacterial; osteopachic; dermatological; antigout;
 XX immunomodulator; antiarrhythmic; cardiac; nootropic; antihaemic;
 XX nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
 XX antidiabetic; anabolic; hypertensive; vulnery; gene therapy; cancer;
 XX reproductive system disorder; therapeutic protein.
 XX
 XX Unidentified.
 XX
 XX PN US2004010134-A1.
 XX
 XX PD 15-JAN-2004.
 XX
 XX PF 12-APR-2001; 2001US-00833245.
 XX
 XX PR 12-APR-2000; 2000US-0229358P.
 XX 25-APR-2000; 2000US-0199384P.
 XX 21-DEC-2000; 2000US-0256931P.
 XX
 XX (ROSE/) ROSEN C A.
 XX (HASE/) HASELTINE W A.
 XX
 XX PI Rosen CA, Haseeltine WA;
 XX WPI; 2004-090519/09.
 XX
 XX New albumin fusion proteins, useful for diagnosing, treating, preventing
 XX or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,
 XX asthma, inflammatory bowel disease or Alzheimer's disease.
 XX
 XX Disclosure; SEQ ID NO 369; 279pp; English.
 XX
 XX The invention relates to a novel albumin fusion protein. The invention
 XX further relates to: a composition comprising the albumin fusion protein
 XX and a pharmaceutical carrier; a kit comprising the composition of the
 XX albumin fusion protein formula; a method of treating a disease or
 XX disorder in a patient comprising the step of administering the albumin
 XX fusion protein; a method of treating a patient with a disease or disorder
 XX that is modulated by Therapeutic protein: X, or its fragment or variant;
 XX a method of extending the shelf life of Therapeutic protein: X, or its
 XX fragment or variant; a nucleic acid molecule comprising a polynucleotide
 XX sequence encoding the albumin fusion protein; a vector comprising the
 XX nucleic acid molecule of the albumin fusion protein; and a host cell
 XX comprising the nucleic acid molecule of the albumin fusion protein. The
 XX albumin fusion protein and its compositions have the following
 XX activities: cytosratic, antanaemic, antiarthritic, antiasthmatic, anti-
 XX HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial,
 XX osteopachic, dermatological, antigout, immunomodulator, antiarrhythmic,
 XX cardiac, nootropic, antihaemic, nephrotropic, uropathic,
 XX neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic,
 XX hypertensive, and vulnery. The albumin fusion protein nucleic acid may
 XX be used in gene therapy to treat disorders. The albumin fusion protein is
 XX useful for diagnosing, treating, preventing or ameliorating diseases or
 XX disorders comprising indication: Y. The diseases or disorders include:
 XX cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer),
 XX immune or haematopoietic diseases (e.g. anaemia, Hodgkin's disease, acute
 XX lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,
 XX autoimmune disease, inflammatory bowel disease, psoriasis or Lyme
 XX disease), reproductive system disorders (e.g. prostaticitis, inguinal
 XX hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-
 XX Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,
 XX Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy
 XX or cachexia), cardiovascular disease (e.g. rhabdomyoma, heart disease,
 XX arrhythmia, cardiac arrest, heart valve disease, hypernatremia or

CC hypogastric, mixed foetal diseases (e.g. foetal alcohol syndrome,
CC Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay
CC Sachs disease), excretory diseases (e.g. urinary incontinence, urinary
CC tract infections or renal disorders), neural or sensory disease (e.g.
CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,
CC cerebellar ataxia, attention deficit disorder, autism or obsessive
CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or
CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's
CC disease or glomerulonephritis), digestive diseases (e.g. portal
CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)
CC or connective tissue or epithelial diseases (e.g. Crohn's disease,
CC scleroderma, wound healing or epidermolysis bullosa). This sequence
CC represents a therapeutic protein X relating to the albumin fusion protein
CC of the invention. The sequence listing data for this specification was
CC downloaded from the USPTO website.

XX Sequence 211 AA;

Alignment Scores:

Pred. No.:	55.4	Length:	211
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.00%	Indels:	0
DB:	8	Gaps:	0

US-10-030-271-1 (1-909) X ADL76887 (1-211)

QY 688 CCAGCTGCGCGCCGCGCTGGGCGC 662

DB 126 ProAlaAlaAlaProAlaProGlyAla 134

RESULT 44

ABP62772

ID ABP62772 standard; protein; 262 AA.

AC ABP62772;

XX 23-OCT-2002 (first entry)

XX Protein fragment #9 of S. roseosporus biosynthetic gene cluster.

XX Daptomycin biosynthetic gene cluster; thioesterase; antibacterial;

XX functional; virulence; antiparasitic; immunomodulatory; antileptic;

XX cytostatic; gene therapy; antimicrobial; immunomodulatory; siderophore;

XX anti-cholesterol; agrochemical.

XX Streptomyces roseosporus.

XX MO200259322-A2.

XX 01-AUG-2002.

XX 17-OCT-2001; 2001WO-US032354.

XX 17-OCT-2000; 2000US-0240879P.

XX 28-FEB-2001; 2001US-0272207P.

XX 06-AUG-2001; 2001US-0310385P.

XX (MIAO/) MIAO V P W.

XX (BRI/) BRIAN P.

XX (BAL/) BALZ R H.

XX (SILV/) SILVA C J.

XX Miao VPW, Brian P, Balz RH, Silva CJ;

XX WPI; 2002-599794/64.

XX Isolated nucleic acid molecule from a bacterial daptomycin biosynthetic

XX gene cluster encoding a thioesterase or thioesterase domain, useful for

XX generating novel linear and cyclic peptides, and products in a cell.

XX Claim 138; Page 174; 227pp; English.

XX The invention relates to a novel isolated nucleic acid molecule
CC comprising a sequence that encodes a thioesterase or thioesterase domain,
CC derived from a bacterial daptomycin biosynthetic gene cluster. The
CC proteins of the invention have antibacterial, fungicide, viricide,
CC antiparasitic, immunomodulatory, antileptic, and cytostatic activity. The
CC polynucleotides may have a use in gene therapy. The compositions and
CC methods of the present invention are useful for generating novel linear
CC and cyclic peptides and improving yield of a product in a cell expressing
CC an daptomycin non-ribosomal peptide synthetase (NRPS) to be used as new
CC compounds or in producing new compounds, such as antibiotics,
CC antifungals, antivirals, antiparasitics, antileptics, antitumour agents,
CC immunomodulatory agents, anti-cholesterol agents, siderophores,
CC agrochemicals and cytostatics. The sequences shown in ABP62764-ABQ62808
CC represent proteins encoded by genes derived from a S. roseosporus
CC daptomycin biosynthetic gene cluster

XX Sequence 262 AA;

Alignment Scores:

Pred. No.:	53.6	Length:	262
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.97%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-1 (1-909) X ABP62772 (1-262)

QY 807 CGCGCTGTCAGAGCCCTGCGGCGCT 833

DB 18 ArgProAlaAlaGlyProAlaGlyArg 26

RESULT 45

ADJ72183

ID ADJ72183 standard; protein; 262 AA.

AC ADJ72183;

XX 06-MAY-2004 (first entry)

XX Streptomyces roseosporus daptomycin biosynthesis gene cluster protein #9.

XX antibacterial; gene therapy; daptomycin biosynthesis gene cluster;

XX daptomycin non-ribosomal peptide synthetase; DptBC;

XX gram-positive bacterial infection.

XX Streptomyces roseosporus.

XX MO2003014297-A2.

XX 31-JUL-2002; 2002WO-US024310.

XX 06-AUG-2001; 2001US-0310385P.

XX 17-OCT-2001; 2001WO-US032354.

XX 10-MAY-2002; 2002US-0379866P.

XX (CUBI-) CUBIST PHARM INC.

XX Miao VPW, Brian P, Balz RH, Coeflet-Jegal MF;

XX WPI; 2003-268192/26.

XX New isolated nucleic acid molecule encoding a daptomycin non-ribosomal

XX peptide synthetase, useful for treatment of a gram-positive bacterial

XX infection of skeletal muscle, skin, bloodstream, kidneys, heart, lung and

XX bone.

XX disclosure, SEQ ID NO 31; 292pp; English.

XX The invention relates to new isolated nucleic acid (NA) molecules from

PS Claim 27; SEQ ID NO 872; 591bp; English.
XX
XX The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithp (ACC6080-ACC46749) and to their encoded
CC protein (DITHP; ABR41136-ABR41812). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithp cDNA
CC sequence of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithp nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
CC detecting dithp nucleotide and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithp hybridisation
CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which has transcription
CC factor activity. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 309 AA;
Alignment Scores:
Pred. No.: 52.3 Length: 309
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: Gaps: 0
US-10-030-271-1 (1-909) x ABR41337 (1-309)
QY 688 CCAAGTCGGCGCGCAGCGCTGGGGCC 662
DB 224 ProAlaAlaProAlaProGlyAla 232
RESULT 48
ID ABO69778 standard; protein; 331 AA.
XX ABO69778;
AC ABO69778;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #1953.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Rubenfield MJ, Nolling J, DeLoughery C, Bush D;
XX
XX MPI; 2003-615309/58.
DR N-PSDB; ABO03349.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 18524; 455bp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 331 AA;
Alignment Scores:
Pred. No.: 51.8 Length: 331
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: Gaps: 0
US-10-030-271-1 (1-909) x ABO69778 (1-331)
QY 657 CCGCGCGCGCCAGAGCGCTGGCGGCA 683
DB 124 ProAlaAlaProGlyAlaGlyAlaAla 132
RESULT 49
ID AAB79137 standard; protein; 336 AA.
XX AAB79137;
AC AAB79137;
XX
XX 30-APR-2001 (first entry)
XX
XX Corynebacterium glutamicum HA protein sequence SEQ ID NO:230.
XX
XX Corynebacterium glutamicum; homeostasis; adaptation; HA protein;
KM fine chemical production; organic acid; proteogenic amino acid;
KM nonproteogenic amino acid; purine base; pyrimidine base; nucleoside;
KM nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KM carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KM diagnosis; Corynebacterium diptheriae; genetic engineering;
KM Brevibacterium; environmental condition.
XX
XX Corynebacterium glutamicum.
XX
XX WO200100842-A2.
XX
XX 04-JAN-2001.
XX
XX 23-JUN-2000; 2000MO-IB000911.
XX
XX 25-JUN-1999; 99US-0141031P.
XX
XX 08-JUL-1999; 99DE-01031636.
XX
XX 09-JUL-1999; 99DE-01032125.
XX

CC production. The C. glutamicum SMP gene can be used in vectors (II) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteogenic or
 CC nonproteogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polypeptide, or an enzyme. The presence of (I) or SMP proteins (II)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localize C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)
 CC
 XX

SO Sequence 336 AA:

Alignment Scores:

Pred. No.:	51.7	Length:	336
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.97%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-1 (1-909) X AAB79487 (1-336)

QY	844	GAGGCCCTGTGGAGGCTGTGGCCCG	870
DB	58	GlutAlaLeuArgGluAlaValGlyArg	66

Search completed: March 22, 2005, 03:00:32
 Job time : 142.322 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 03:04:23 ; Search time 240.761 Seconds
(without alignments)
2495.932 Million cell updates/sec

Title: US-10-030-271-1
Perfect score: 303
Sequence: 1 atgcgcctaccggtcgcac.....tcagtgtgatagcgtgcac 909

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1401741 seqs, 330541175 residues

Word size: 1

Total number of hits satisfying chosen parameters: 268658

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame_n2p model -DEV=x1p
-Q=/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
-DB=Published Applications AA -QFMT=fastan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPCH=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human4.cdi -LIST=100 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=us10030271 @CGN 1.1 418 @runat 21032005.15550-26862
-NCU=6 -ICPU=3 -NO MMAR -LARGESUBSET -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09D_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match	Length	DB	ID	Description
1	224	73.9	304	14	US-10-013-477-11	Sequence 11, Appl
2	224	73.9	318	14	US-10-001-254-18	Sequence 18, Appl
3	224	73.9	326	14	US-10-296-539-1	Sequence 1, Appl
4	224	73.9	366	14	US-10-106-698-4626	Sequence 4626, Ap
5	194	64.0	319	9	US-09-925-302-758	Sequence 758, App
6	194	64.0	319	10	US-09-925-302-758	Sequence 758, App
7	134	44.2	242	15	US-10-296-115-1220	Sequence 1220, Ap
8	107	35.3	217	9	US-09-799-777-26	Sequence 26, Appl
9	67	22.1	101	14	US-10-001-254-8	Sequence 8, Appl
10	67	22.1	101	14	US-09-733-167-6	Sequence 6, Appl
11	13	4.3	210	9	US-09-935-823-4	Sequence 4, Appl
12	13	4.3	318	9	US-09-935-823-6	Sequence 6, Appl
13	13	4.3	318	9	US-09-935-823-6	Sequence 1, Appl
14	13	4.3	318	9	US-09-733-167-1	Sequence 3, Appl
15	13	4.3	318	9	US-09-733-167-3	Sequence 10755,
16	11	3.7	107	16	US-10-437-963-107955	Sequence 166242,
17	11	3.6	145	16	US-10-437-963-166242	Sequence 38, Appl
18	10	3.3	146	14	US-10-001-254-38	Sequence 23101,
19	9	3.0	66	15	US-10-424-599-231701	Sequence 47327, A
20	9	3.0	90	16	US-10-767-701-47327	Sequence 55826, A
21	9	3.0	94	16	US-10-767-701-55826	Sequence 255144,
22	9	3.0	115	15	US-10-424-599-255144	Sequence 48342, A
23	9	3.0	116	16	US-10-767-701-48342	Sequence 13396,
24	9	3.0	128	16	US-10-437-963-133296	Sequence 142899,
25	9	3.0	131	16	US-10-437-963-142899	Sequence 164498,
26	9	3.0	133	16	US-10-437-963-164498	Sequence 369, App
27	9	3.0	211	11	US-09-833-245-369	Sequence 56323, A
28	9	3.0	216	15	US-10-425-114-56323	Sequence 197739,
29	9	3.0	262	17	US-10-211-028-31	Sequence 60272, A
30	9	3.0	298	15	US-10-425-114-60272	Sequence 104157,
31	9	3.0	330	16	US-10-437-963-104157	Sequence 490, App
32	9	3.0	336	16	US-10-781-014-490	Sequence 1, Appl
33	9	3.0	370	14	US-10-285-976-1	Sequence 50, Appl
34	9	3.0	370	15	US-10-302-812-50	Sequence 2, Appl
35	9	3.0	370	17	US-10-847-972-2	Sequence 6644, Ap
36	9	3.0	395	9	US-09-738-626-6644	Sequence 13365, A
37	9	3.0	398	14	US-10-156-761-13365	Sequence 20883, A
38	9	3.0	476	15	US-10-369-499-20883	Sequence 155657,
39	9	3.0	477	16	US-10-437-963-155657	Sequence 10964, A
40	9	3.0	478	14	US-10-156-761-10964	Sequence 36, Appl
41	9	3.0	503	8	US-08-459-455-36	Sequence 8, Appl
42	9	3.0	503	14	US-10-123-529-8	Sequence 5, Appl
43	9	3.0	505	9	US-09-888-243-5	Sequence 62674, A
44	9	3.0	517	15	US-10-425-114-62674	Sequence 42966, A
45	9	3.0	564	15	US-10-425-114-42966	Sequence 149484,
46	9	3.0	675	16	US-10-437-963-149484	Sequence 142715,
47	9	3.0	901	16	US-10-437-963-142715	Sequence 1972, Ap
48	9	3.0	955	16	US-10-408-765A-1972	Sequence 2346, Ap
49	9	3.0	962	15	US-10-276-774-2346	Sequence 62666, A
50	9	3.0	2031	15	US-10-282-122A-62666	Sequence 1528, Ap
51	9	2.6	10	10	US-09-572-404B-1528	Sequence 1530, Ap
52	8	2.6	10	10	US-09-572-404B-1530	Sequence 150, App
53	8	2.6	15	9	US-09-879-957-150	Sequence 94, Appl
54	8	2.6	15	14	US-10-185-050-94	Sequence 150, App
55	8	2.6	15	14	US-10-161-791-307	Sequence 35, Appl
56	8	2.6	15	16	US-10-807-865-150	Sequence 23, Appl
57	8	2.6	21	14	US-10-116-252-35	Sequence 31, Appl
58	8	2.7	27	9	US-09-896-888A-29	Sequence 186712,
59	8	2.7	28	9	US-09-896-888A-29	Sequence 27, Appl
60	8	2.6	43	9	US-09-896-888A-27	Sequence 102706,
61	8	2.6	43	16	US-10-437-963-102706	Sequence 106496,
62	8	2.6	51	16	US-10-437-963-106496	Sequence 189336,
63	8	2.6	60	16	US-10-437-963-189336	Sequence 156496,
64	8	2.6	61	15	US-10-424-599-156496	Sequence 23491, A
65	8	2.7	62	14	US-10-029-386-32491	Sequence 216648,
66	8	2.7	70	15	US-10-424-599-216648	Sequence 195679,
67	8	2.7	71	15	US-10-424-599-195679	Sequence 255309,
68	8	2.7	71	15	US-10-424-599-255309	Sequence 220813,
69	8	2.7	71	15	US-10-424-599-220813	Sequence 237140,
70	8	2.7	73	15	US-10-424-599-237140	
71	8	2.6	76	15	US-10-424-599-237140	

```
c 72 8 2.7 78 16 US-10-767-701-58510 Sequence 58510, A
73 8 2.6 79 16 US-10-437-963-174687 Sequence 174687,
74 8 2.6 80 16 US-10-437-963-129829 Sequence 129829,
75 8 2.6 80 16 US-10-437-963-150240 Sequence 150240,
76 8 2.6 82 16 US-10-437-963-198587 Sequence 198587,
77 8 2.6 85 15 US-10-767-701-49110 Sequence 49110, A
78 8 2.7 86 15 US-10-425-114-48694 Sequence 48694, A
79 8 2.6 87 16 US-10-437-963-127599 Sequence 127599,
c 80 8 2.7 91 16 US-10-437-963-201366 Sequence 201366,
81 8 2.6 93 16 US-10-437-963-177858 Sequence 177858,
c 82 8 2.7 94 16 US-10-437-963-131432 Sequence 131432,
83 8 2.6 97 16 US-10-437-963-105618 Sequence 105618,
c 84 8 2.7 97 16 US-10-437-963-192695 Sequence 192695,
85 8 2.7 99 16 US-10-767-701-61057 Sequence 61057, A
c 86 8 2.6 103 16 US-10-437-963-112949 Sequence 112949,
c 87 8 2.7 104 15 US-10-424-599-241061 Sequence 241061,
c 88 8 2.6 106 16 US-10-437-963-147002 Sequence 147002,
c 89 8 2.7 107 16 US-10-437-963-141465 Sequence 141465,
c 90 8 2.7 108 16 US-10-437-963-124659 Sequence 124659,
91 8 2.6 111 15 US-10-296-115-825 Sequence 825, App
92 8 2.6 113 16 US-10-437-963-140160 Sequence 140160,
93 8 2.6 114 16 US-10-437-963-144007 Sequence 144007,
c 94 8 2.7 120 16 US-10-437-963-147321 Sequence 147321,
c 95 8 2.6 120 16 US-10-060-036-68 Sequence 61593, A
96 8 2.6 125 14 US-10-060-036-68 Sequence 68, App
97 8 2.6 126 16 US-10-437-963-152734 Sequence 152734,
c 98 8 2.7 129 15 US-10-424-599-267849 Sequence 267849,
c 99 8 2.6 129 15 US-10-276-774-1615 Sequence 1615, Ap
c 100 8 2.7 137 16 US-10-437-963-141132 Sequence 141132,
```

ALIGNMENTS

```
RESULT 1
; Sequence 11, Application US/10013477
; Publication No. US20030049732A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Apoptosis Related Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO02P1
; CURRENT APPLICATION NUMBER: US/10/013,477
; CURRENT FILING DATE: 2001-12-13
; PRIOR APPLICATION NUMBER: 09/669,445
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: PCT/US00/06642
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/126,018
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/139,638
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/149,449
; PRIOR FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-013-477-11

Alignment Scores:
Pred. No.: 2,21e-182 Length: 304
Score: 224.00 Matches: 224
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.93% Indels: 0
Gaps: 0
DB: 14

US-10-030-271-1 (1-909) x US-10-013-477-11 (1-304)
```

```
QY 238 GAAGCACTGGCTGCTGGGCAACTCTGCGCTGCTGGCCGCAACACTGCTG
|||||
```

```
Db 58 GluSerAsnLeuAArgLeuLeuGlyGlnLeuLeuAArgValLeuAlaArgHisAspLeuLeu 77
298 CCGCACTTGCGCGCAAGCGCGCGCGCAAGTGTCTCCGAAACGCTATAGTATGGCAC 357
Db 78 ProHisLeuAlaAArgValArgArgArgProValSerProGluAArgTysSerIArgIThr 97
QY 358 TCAGGCTCTTCAAGAGACAGAGGGTACCTGCCGTCCGCGTCCGACAGTCAAGCACTTCT 417
98 SerSerSerSerIArgThrIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgI 117
QY 418 GGAATTCACAGAGGCTAGTGGAGACAGAGCTCCCGCCCAACCAAGGCGAGCGCGCG 477
118 AlaAsnSerGlnGlnGlnIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgI 137
QY 478 ACTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
Db 138 SerArgIArgIArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaPro 157
QY 538 CAGCAGCACTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
158 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArg 177
QY 598 CTCGGGCTTGAAGAGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 657
178 LeuArgValAArgAlaGlnIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgI 197
Db 658 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 717
198 ArgArgProGlnAlaLeuAlaArgGlnLeuAlaPheGlyGlnAlaThrAlaValLeu 217
QY 718 CGCTCAAGGAGACTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 777
218 ArgSerArgAspLeuGlySerValValCysAspIleValPheSerGlnLeuSerIArgIArg 237
QY 778 GAGCGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 837
238 AspAlaPheArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArgIArg 257
Db 838 CTGACTGAGCGCGCTGCGAGAGGCTGTGGCGCGGAGGCTGTTCGCTGTGTGTGTGTGTG 897
258 LeuThrGluAlaLeuArgGlnAlaValGlyArgGlnAlaValAArgLeuLeuValSerVal 277
QY 898 GATGAGGCTGAC 909
Db 278 AspGluAlaAsp 281
```

```
RESULT 2
; Sequence 18, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfred
; APPLICANT: Stenner-Liewen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-001-254-18
```



```
Db 300 AspgluAlaasp 303
RESULT 4
US-10-106-698-4626
; Sequence 4626, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruden et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4626
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-4626

Alignment Scores:
Pred. No.: 2,12e-182 Length: 366
Score: 224.00 Matches: 224
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.93% Indels: 0
DB: 14 Gaps: 0

US-10-030-271-1 (1-909) x US-10-106-698-4626 (1-366)
QY 238 GAGAGCACTCTGGGCTGCTGGGGCAACTCTCCGCGTGCTGGCCCGCCAGCACTGCTG 297
Db 120 GluSerAsnLeuArgLeuArgLeuGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 139
QY 298 CCGCACTCTGGCGGCAAGCGCGCCGCGCAAGTGTCTCCAGAAAGCTATAGCTATGGACCC 357
Db 140 ProHisLeuAlaArgHisArgValArgProValSerProGluArgTyrSerTyrGlnTyr 159
QY 358 TCAGGCTCTTCAAAAGAGACAGAGGGTAGCTGCTGCGCGTGCGGCAAGTCAAGCACTTCT 417
Db 160 SerSerSerSerTyrArgThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179
QY 418 GCAAAATTCACAGAGGTCAGTGGGAGACAGGCTCCCCCCCCCAACCAAGCGGCGCGCG 477
Db 180 AlaAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
QY 478 AGTCGGGGCGCGCCAGTGGTGTGCACAGCGCGCGGAGAGGGGCCCGCCAGCGCGACCC 537
Db 200 SerArgGlnArgProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 219
QY 538 CAGCAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAGTGAAGTGAAGTGAAGTCCG 597
Db 220 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGlnGln 239
QY 598 CTCGGGGTTTGACAGAGTACTCTGCGAGCATGGGCGCAGCTTGGAGCAGGCGCTGCATCC 657
Db 240 LeuArgValArgAlaGlnTyrCysGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 259
QY 658 CCGCGGCGCCCGCGGCGTGGCGCGGCGAGCTGAGTGTGGGAGGCGCAGCGAGTGTG 717
Db 260 ArgArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlnGlnAlaThrAlaValLeu 279
QY 718 CGGTCAAGGAGCCTGGGCTCTGTGTTTGACATCAAGTTCACAGGCTTCTCTATCTG 777
Db 280 ArgSerArgAspLeuGlnSerValValCysAspIleLysPheSerGlnLeuSerTyrLeu 299
QY 778 GACGCTTCTGGGCGAGTACTGAGTGGCGCGCTGCTGCGAGGCGCTTGGCGGCGGCTTC 837
```

```
Db 300 AspaAlaPheTyrGlnAspTyrLeuSerGlnValAlaLeuGlnAlaLeuArgGlnValPhe 319
QY 838 CTGACTGAGGCCCTCGAGAGGCTGTGGCCCGGAGAGGCTGTTCGCTGCTGCTAGTGTG 897
Db 320 LeuThrGlnAlaLeuArgGlnAlaValGlnArgGlnAlaValArgLeuLeuValSerVal 339
QY 898 GATGAGGCTGAC 909
Db 340 AspgluAlaasp 343

RESULT 5
US-09-925-302-758
; Sequence 758, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1998-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-758

Alignment Scores:
Pred. No.: 9,16e-157 Length: 319
Score: 194.00 Matches: 194
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 64.03% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-925-302-758 (1-319)
QY 238 GAGAGCACTCTGGGCTGCTGGGGCAACTCTCCGCGTGCTGGCCCGCCAGCACTGCTG 297
Db 120 GluSerAsnLeuArgLeuArgLeuGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 139
QY 298 CCGCACTCTGGCGGCAAGCGCGCCGCGCAAGTGTCTCCAGAAAGCTATAGCTATGGACCC 357
Db 140 ProHisLeuAlaArgHisArgValArgProValSerProGluArgTyrSerTyrGlnTyr 159
QY 358 TCAGGCTCTTCAAAAGAGACAGAGGGTAGCTGCTGCGCGTGCGGCAAGTCAAGCACTTCT 417
Db 160 SerSerSerSerTyrArgThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179
QY 418 GCAAAATTCACAGAGGTCAGTGGGAGACAGGCTCCCCCCCCCAACCAAGCGGCGCGCG 477
Db 180 AlaAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
QY 478 AGTCGGGGCGCGCCAGTGGTGTGCACAGCGCGCGGAGAGGGGCCCGCCAGCGCGACCC 537
Db 200 SerArgGlnArgProSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 219
QY 538 CAGCAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAGTGAAGTGAAGTGAAGTCCG 597
Db 220 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGlnGln 239
QY 598 CTCGGGGTTTGACAGAGTACTCTGCGAGCATGGGCGCAGCTTGGAGCAGGCGCTGCATCC 657
Db 240 LeuArgValArgAlaGlnTyrCysGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 259
QY 658 CCGCGGCGCCCGCGGCGTGGCGCGGCGAGCTGAGTGTGGGAGGCGCAGCGAGTGTG 717
```

```
Db      260 ArgArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeu 279
Qy      718 CGCTCAAGGAGGACCTGGGCTCTGTGTTTGACATCAAGTTCTCAGAGCTTCTCTATCTG 777
        |||
Db      280 ArgSerArgAspLeuGlySerValValCysAspIleValPheSerGlnLeuSerTyrlu 299
Qy      778 GACGCTTCTGGGGCGACTACTAGTGGCGCCCTGCTGCAG 819
        |||
Db      300 AspAlaPheTrpGlyAspTyrluSerGlyAlaLeuLeuGln 313

RESULT 6
US-09-925-302-758
; Sequence 758, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-758

Alignment Scores:
Pred. No.:      9 16e-157      Length:      319
Score:          194.00      Matches:      194
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    64.03%      Indels:      0
DB:             10      Gaps:      0

US-10-030-271-1 (1-909) x US-09-925-302-758 (1-319)
Qy      238 GAGAGCAACCTGGGCTGTGCTGGGCAACTCTGGCGCTGTGCGCCGCAAGACTGCTG 297
        |||
Db      120 GlnSerSerLeuArgLeuLeuGlnGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 139
Qy      298 CCGGCACTGGCGCGCAAGCGCGCGCGCGCTCTCCAGAACGCTATAGCTATGCGACC 357
        |||
Db      140 ProHisLeuAlaArgIleArgValArgArgProValSerProGlnArgTyrluSerTyrluThr 159
Qy      358 TCCAGCTCTTCAAGAGAGAGAGAGAGGAGTGGCGCTGGCGCGCTGAGTCAAGAGATTCT 417
        |||
Db      160 SerSerSerSerLeuArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179
Qy      418 GCAAAATTCAGACAGGCTCAGTGGAGAGAGAGCTCCCGCCCAACCAAGCGGACGCGCG 477
        |||
Db      180 AlaAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
Qy      478 AGTCGGGGCGCGCGCGCTGCTGCTGCGCAAGCGCGCGCGAGAGGGCGCCCAAGCGCGACC 537
        |||
Db      200 SerArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaPro 219
Qy      538 CAGAGAGAGTCAAGCGCGCGCGCAAGCTTCTCTGAAGGCAAGGAGACTGTGACATCCG 597
        |||
Db      220 GlnGlnGlnSerGlnProAlaArgProSerSerIleGlyValIleThrCysAspIleArg 239
Qy      598 CTCGGGTTTCAGACAGAGTACTGCGACATAGGCGCGCAAGCTTGAAGAGCGCGTGGCATCC 657
        |||
Db      240 LeuArgValArgIleArgIleGlyArgGlnHisGlyProAlaLeuGlnGlnGlnGlnGln 259
Qy      658 CGGCGCGCGCGCGCGCTGGCGCGCGCGCGAGCTGACCTGTTGGGCGCGCGCAAGCGCGCTG 717
        |||
Db      260 ArgArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeu 279
```

```
Qy      718 CGCTCAAGGAGGACCTGGGCTCTGTGTTTGACATCAAGTTCTCAGAGCTTCTCTATCTG 777
        |||
Db      280 ArgSerArgAspLeuGlySerValValCysAspIleValPheSerGlnLeuSerTyrlu 299
Qy      778 GACGCTTCTGGGGCGACTACTAGTGGCGCCCTGCTGCAG 819
        |||
Db      300 AspAlaPheTrpGlyAspTyrluSerGlyAlaLeuLeuGln 313

RESULT 7
US-10-296-115-1220
; Sequence 1220, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1220
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1220

Alignment Scores:
Pred. No.:      1 71e-105      Length:      242
Score:          134.00      Matches:      134
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    44.22%      Indels:      0
DB:             15      Gaps:      0

US-10-030-271-1 (1-909) x US-10-296-115-1220 (1-242)
Qy      508 CGGCGGCGAGAGGGGCGCCGCGCAAGCCGACAGAGTCAAGAGCCCGCGACCTTCC 567
        |||
Db      86 ArgArgArgArgGlyAlaProAlaAlaProGlnGlnGlnSerGlnProAlaArgProSer 105
Qy      568 TCTGAAGGCAAAATGACCTGTGACATCCGCGCTCGGGTTGAGCAGAGTACGCGAGCAT 627
        |||
Db      106 SerGlnGlyIleValIleThrCysAspIleArgLeuArgValArgIleGlyIleHis 125
Qy      628 GGGCCAGCTTGGAGAGAGGCGTGGCATCCCGCGCGCGCCGAGCGGCTGGCGCGGAGCTG 687
        |||
Db      126 GlyProAlaLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 145
Qy      688 GACGTGTTGGGAGGCGCACCGGAGTGGCGCTCAAGGAGACTGGGCTGTGTGTTGT 747
        |||
Db      146 AspValPheGlyGlnAlaThrAlaValAlaValLeuArgSerArgAspLeuGlySerValCys 165
Qy      748 GATATCAAGTTCTCAGAGCTTCTCTATCTGAGAGCTTCTGGGGCGAGTCAAGTGGC 807
        |||
Db      166 AspIleValPheSerSerGlnLeuSerTyrluAspAlaPheTrpIleAspTyrluSerGly 185
Qy      808 GCCCTCTGCAAGGCGCTGGCGGGCGGTTCCTGACTGAGGCGCTTGGAGAGGCTGTGGGC 867
        |||
Db      186 AlaLeuLeuGlnAlaLeuArgGlyValAlaPheLeuThrGlnAlaLeuArgGlnAlaValGly 205
Qy      868 CGGAGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
        |||
Db      206 ArgGlnAlaValArgLeuLeuValSerValAspGlnAlaAsp 219

RESULT 8
US-09-799-777-26
; Sequence 26, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
```


;; PRIOR APPLICATION NUMBER: PCT/DE99/01712
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: German Patent Application No. US20020099009A1 198 25 621.3
;; PRIOR FILING DATE: 1998-06-08
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 6
;; LENGTH: 210
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: C-DEDD, which is a deletion mutant of human DEDD comprising amino
;; OTHER INFORMATION: acids 109-318 of the naturally occurring human DEDD.
US-09-733-167-6

Alignment Scores:
Pred. No.: 0.038 Length: 210
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-733-167-6 (1-210)

Qy 583 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 621
Db 84 ThrcysaplleargleuargValargAlaGlutryCys 96

RESULT 11
US-09-935-223-4
;; Sequence 4, Application US/09935223
;; Publication No. US20020086983A1
;; GENERAL INFORMATION:
;; APPLICANT: Alnemri, Emad S.
;; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, And
;; FILE REFERENCE: T02499
;; CURRENT APPLICATION NUMBER: US/09/935,223
;; CURRENT FILING DATE: 2001-08-22
;; PRIOR APPLICATION NUMBER: 09/723,450
;; PRIOR FILING DATE: 2000-11-28
;; PRIOR APPLICATION NUMBER: 09/276,993
;; PRIOR FILING DATE: 1999-03-26
;; PRIOR APPLICATION NUMBER: 08/859,167
;; PRIOR FILING DATE: 1997-05-20
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 4
;; LENGTH: 318
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Novel Sequence
US-09-935-223-4

Alignment Scores:
Pred. No.: 0.0356 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-935-223-4 (1-318)

Qy 583 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 621
Db 192 ThrcysaplleargleuargValargAlaGlutryCys 204

RESULT 12
US-09-935-223-6
;; Sequence 6, Application US/09935223

;; Publication No. US20020086983A1
;; GENERAL INFORMATION:
;; APPLICANT: Alnemri, Emad S.
;; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, And
;; FILE REFERENCE: T02499
;; CURRENT APPLICATION NUMBER: US/09/935,223
;; CURRENT FILING DATE: 2001-08-22
;; PRIOR APPLICATION NUMBER: 09/723,450
;; PRIOR FILING DATE: 2000-11-28
;; PRIOR APPLICATION NUMBER: 09/276,993
;; PRIOR FILING DATE: 1999-03-26
;; PRIOR APPLICATION NUMBER: 08/859,167
;; PRIOR FILING DATE: 1997-05-20
;; NUMBER OF SEQ ID NOS: 17
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 6
;; LENGTH: 318
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Novel Sequence
US-09-935-223-6

Alignment Scores:
Pred. No.: 0.0356 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-935-223-6 (1-318)

Qy 583 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 621
Db 192 ThrcysaplleargleuargValargAlaGlutryCys 204

RESULT 13
US-09-733-167-1
;; Sequence 1, Application US/09733167
;; Patent No. US20020099009A1
;; GENERAL INFORMATION:
;; APPLICANT: Peter, Marcus
;; TITLE OF INVENTION: Protein for Regulation of Apoptosis
;; FILE REFERENCE: 4121-120
;; CURRENT APPLICATION NUMBER: US/09/733,167
;; CURRENT FILING DATE: 2000-12-08
;; PRIOR APPLICATION NUMBER: PCT/DE99/01712
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: German Patent Application No. US20020099009A1 198 25 621.3
;; PRIOR FILING DATE: 1998-06-08
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 318
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-733-167-1

Alignment Scores:
Pred. No.: 0.0356 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-733-167-1 (1-318)

Qy 583 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 621
|||||

```

DB      192 ThCysAspIleArgLeuArgValArgAlaGluTrpCys 204

RESULT 14
US-09-733-167-3
; Sequence 3, Application US/09733167
; Patent No. US20020099009A1
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcue
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US20020099009A1 198 25 621.3
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-167-3

Alignment Scores:
Pred. No.: 0.0356 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.29% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-733-167-3 (1-318)

QY      583 ACCTGTGACATCGGCTCGGGTTTCGAGCAGACTGTC 621
DB      192 ThCysAspIleArgLeuArgValArgAlaGluTrpCys 204

RESULT 15
US-10-437-963-107955
; Sequence 107955, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 107955
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12256C.1.pep
US-10-437-963-107955

Alignment Scores:
Pred. No.: 2.28 Length: 107
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.67% Indels: 0
DB: 16 Gaps: 0

```

```

US-10-030-271-1 (1-909) x US-10-437-963-107955 (1-107)

QY      265 GTTGCCCGACAGCCGCGAGTTGCTCTCCGCC 233
DB      17 ValAlaProAlaAlaAlaGlyCysSerArgArg 27

RESULT 16
US-10-437-963-166242
; Sequence 166242, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 166242
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64970C.1.pep
US-10-437-963-166242

Alignment Scores:
Pred. No.: 2.14 Length: 145
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.63% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-166242 (1-145)

QY      505 AGACGGCGCGAGAGAGGCGCCCGCAGCCG 537
DB      51 ArgArgArgArgGlyAlaProAlaAlaPro 61

RESULT 17
US-10-001-254-38
; Sequence 38, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfred
; APPLICANT: Steiner-liewen, Frank
; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
; FILE REFERENCE: P-Lt 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 146
; TYPE: PRT

```

ORGANISM: Danio rerio
US-10-001-254-38

Alignment Scores:

Pred. No.:	Length:
Score: 15.2	Matches: 146
Percent Similarity: 10.00	Conservative: 10
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 100.00%	Indels: 0
DB: 14	Gaps: 0

US-10-030-271-1 (1-909) x US-10-001-254-38 (1-146)

QY 199 GAGCTCCTGCTGAGCTGAGCCGCCGCG 228
DB 101 GluLeuLeuGluLeuGluArgArgGly 110

RESULT 18

US-10-424-599-231701
Sequence 231701, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 231701
LENGTH: 66
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_51249C.1.pep
US-10-424-599-231701

Alignment Scores:

Pred. No.:	Length:
Score: 128	Matches: 66
Percent Similarity: 9.00	Conservative: 9
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 100.00%	Indels: 0
DB: 15	Gaps: 0

US-10-030-271-1 (1-909) x US-10-424-599-231701 (1-66)

QY 221 CGCTCCAGCTCCAGCAGAGCTAGG 195
DB 4 ArgSerSerSerArgSerArg 12

RESULT 19

US-10-767-701-47327
Sequence 47327, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic David K.
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 47327
LENGTH: 90
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:

OTHER INFORMATION: Clone ID: SORBI-28MAY03-C98_4.pep
US-10-767-701-47327

Alignment Scores:

Pred. No.:	Length:
Score: 120	Matches: 90
Percent Similarity: 9.00	Conservative: 9
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 100.00%	Indels: 0
DB: 16	Gaps: 0

US-10-030-271-1 (1-909) x US-10-767-701-47327 (1-90)

QY 505 AGACGGCGCGGAGAGGCGCCAGCC 531
DB 47 ArgArgArgArgArgGlyAlaProAla 55

RESULT 20

US-10-767-701-55826
Sequence 55826, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic David K.
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 55826
LENGTH: 94
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 30933130.pep
US-10-767-701-55826

Alignment Scores:

Pred. No.:	Length:
Score: 119	Matches: 94
Percent Similarity: 9.00	Conservative: 9
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 100.00%	Indels: 0
DB: 16	Gaps: 0

US-10-030-271-1 (1-909) x US-10-767-701-55826 (1-94)

QY 178 GCCCGGCGCGGAGAGCTAGAGCTC 204
DB 39 AlaArgAlaArgSerGlyLeuGluLeu 47

RESULT 21

US-10-424-599-255144
Sequence 255144, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 255144
LENGTH: 115
TYPE: PRT
ORGANISM: Glycine max
FEATURE:

```
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72417C.1.pep
US-10-424-599-255144

Alignment Scores:
Pred. No.: 114 Length: 115
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-1 (1-909) x US-10-424-599-255144 (1-115)

QY 524 GCCCTCTCCGCCGCTGTGCACCA 498
DB 45 AAlaProLeuArgArgArgLeuAlaPro 53

RESULT 22
US-10-767-701-48342
; Sequence 48342, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 48342
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3481-027-Pl-K1-B5.pep
US-10-767-701-48342

Alignment Scores:
Pred. No.: 114 Length: 116
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-767-701-48342 (1-116)

QY 221 CGCTCAGCTCCAGCAGAGCTTAGG 195
DB 61 ArgSerSerSerSerArgSerSerArg 69

RESULT 23
US-10-437-963-133296
; Sequence 133296, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133296
```

```
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35180C.1.pep
US-10-437-963-133296

Alignment Scores:
Pred. No.: 112 Length: 128
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-133296 (1-128)

QY 511 CGCGGAGGAGGGCGCCGACGCC 537
DB 2 ArgArgArgGlyAlaProAlaAlaPro 10

RESULT 24
US-10-437-963-142899
; Sequence 142899, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142899
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43860C.1.pep
US-10-437-963-142899

Alignment Scores:
Pred. No.: 111 Length: 131
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-142899 (1-131)

QY 243 CAACCTGCGGCTGTGGGCACTCCT 269
DB 48 GlnProAlaAlaAlaGlyAlaAlaThrPro 56

RESULT 25
US-10-437-963-164498
; Sequence 164498, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
```


APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 164498
LENGTH: 133
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_63391C.1.pep
US-10-437-963-164498

Alignment Scores:
Pred. No.: 111 Length: 133
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-164498 (1-133)

QY 669 GCGCGTGGCGGCGAGCTGAGCTGTT 695

Db 30 G|Y|A|G|Y|A|A|A|A|A|G|Y|A|G|Y|A| 38

RESULT 26
US-09-833-245-369
Sequence 369, Application US/09833245
Publication No. US20040010134A1
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc.
TITLE OF INVENTION: Albumin Fusion Proteins
FILE REFERENCE: PF546PCT
CURRENT APPLICATION NUMBER: US/09/833,245
CURRENT FILING DATE: 2001-04-12
PRIOR APPLICATION NUMBER: 60/229,358
PRIOR FILING DATE: 2000-04-12
PRIOR APPLICATION NUMBER: 60/256,931
PRIOR FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: 60/199,384
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 2267
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 369
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (64)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (113)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-369

Alignment Scores:
Pred. No.: 101 Length: 211
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 11 Gaps: 0

US-10-030-271-1 (1-909) x US-09-833-245-369 (1-211)

QY 688 CCAAGCTCCGCGCGAGCGCTGGGCGC 682

|||||

Db 126 ProAlaAlaProAlaProGlyAla 134

RESULT 27
US-10-425-114-56323
Sequence 56323, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovacic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack B.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 56323
LENGTH: 216
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73131G05_FLI.pep
US-10-425-114-56323

Alignment Scores:
Pred. No.: 101 Length: 216
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-1 (1-909) x US-10-425-114-56323 (1-216)

QY 661 CCGCGCCGAGGCGCTGGCGCGAGCTG 687

Db 164 ArgProGlnAlaLeuAlaArgGlnLeu 172

RESULT 28
US-10-437-963-197739
Sequence 197739, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 197739
LENGTH: 219
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_93469C.1.pep
US-10-437-963-197739

Alignment Scores:
Pred. No.: 100 Length: 219
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0
US-10-030-271-1 (1-909) x US-10-437-963-197739 (1-219)
QY 663 GCGGCGGCGGCGGCGGCGGCGTGA 689
DB 3 AAlPrGlyAlaGlyAlaAlaAlaGly 11
RESULT 29
US-10-211-028-31
; Sequence 31, Application US/10211028
; Publication No. US20050027113A1
; GENERAL INFORMATION:
; APPLICANT: CUBIST PHARMACEUTICALS, INC.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO THE DAPTOMYCIN
; FILE REFERENCE: CUB-12 PCT CIP
; CURRENT APPLICATION NUMBER: US/10/211, 028
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: PCT/US02/24310
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: PCT/US01/32354
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/310,385
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 60/379,866
; PRIOR FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Streptomyces roseosporus
US-10-211-028-31
Alignment Scores:
Pred. No.: 96.6 Length: 262
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 17 Gaps: 0
US-10-030-271-1 (1-909) x US-10-211-028-31 (1-262)
QY 807 CGCCGCTGCTGAGCGCCCTGGCGGCGGT 833
DB 18 ArgPrGlyAlaGlyProAlaGlyArg 26
RESULT 30
US-10-425-114-60272
; Sequence 60272, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425, 114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 60272
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3600-053-C3_FLI.pep

US-10-425-114-60272
Alignment Scores:
Pred. No.: 94.1 Length: 298
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 15 Gaps: 0
US-10-030-271-1 (1-909) x US-10-425-114-60272 (1-298)
QY 499 GCGGCGGCGGCGGCGGCGGCGGCC 525
DB 74 GlyAlaArgArgArgArgArgGlyAla 82
RESULT 31
US-10-437-963-104157
; Sequence 104157, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 104157
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(330)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_101517C.1.pep
US-10-437-963-104157
Alignment Scores:
Pred. No.: 92.1 Length: 330
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0
US-10-030-271-1 (1-909) x US-10-437-963-104157 (1-330)
QY 669 GCGGCTGGCGGCGGCGGCGGCGGCTGT 695
DB 177 GlyAlaGlyAlaAlaAlaGlyArgVal 185
RESULT 32
US-10-781-014-490
; Sequence 490, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Heberhauser, Gregor
; TITLE OF INVENTION: CORNEACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY


```
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 17 Gaps: 0

US-10-030-271-1 (1-909) x US-10-847-972-2 (1-370)

QY 243 GCTCTGCGCGGACTGCGCGGCGCTC 217
DB 18 AAlaLeuAlaAlaLeuProAlaAlaLeu 26

RESULT 36
US-09-738-626-6644
; Sequence 6644, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MAKIKO
; APPLICANT: SENOH, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6644
; LENGTH: 395
; TYPE: PR
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6644

Alignment Scores:
Pred. No.: 88.8 Length: 395
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-1 (1-909) x US-09-738-626-6644 (1-395)

QY 844 GAGGCCCTCGAGAGGCTGTGGCCCG 870
DB 117 GAlaLeuArgGAlaAlaValGlyArg 125

RESULT 37
US-10-156-761-13365
; Sequence 13365, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
```

```
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13365
; LENGTH: 398
; TYPE: PR
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13365

Alignment Scores:
Pred. No.: 88.6 Length: 398
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 14 Gaps: 0

US-10-030-271-1 (1-909) x US-10-156-761-13365 (1-398)

QY 660 GCGGCCCGAGGCGGCTGCGGCGAGCT 686
DB 149 AAlaAlaProGlyAlaGlyAlaAlaAla 157

RESULT 38
US-10-369-493-20883
; Sequence 20883, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20883
; LENGTH: 456
; TYPE: PR
; ORGANISM: SYNECHOCOCCUS SP. WH 8102
US-10-369-493-20883

Alignment Scores:
Pred. No.: 86.2 Length: 456
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-1 (1-909) x US-10-369-493-20883 (1-456)

QY 225 GCGGCGCTCCAGCTCCAGCAGAGACTC 199
DB 149 AAlaAlaLeuGlnLeuGlnGlnGlnLeu 157

RESULT 39
US-10-437-963-155657
; Sequence 155657, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 155657
LENGTH: 477
TYPE: PRN
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_553C.1.pep
US-10-437-963-155657

Alignment Scores:
Pred. No.: 85.4 Length: 477
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-155657 (1-477)

QY 301 GCGGACAGAGTGTGCGGGCCAGCA 275
DB 6 AAlaAlaGlyATGtGtGtGtGtGtA 14

RESULT 40
US-10-156-761-10964
Sequence 10964, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10964
LENGTH: 478
TYPE: PRN
ORGANISM: Streptomyces avermectilis
US-10-156-761-10964

Alignment Scores:
Pred. No.: 85.3 Length: 478
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 14 Gaps: 0

US-10-030-271-1 (1-909) x US-10-156-761-10964 (1-478)

QY 838 CTGACTGAGCCCTGCGAGAGGCTGTG 864
DB 25 LeuThrGlnAlaLeuArgGlnAlaVal 33

RESULT 41

US-08-459-455-36
Sequence 36, Application US/08459455
Publication No. US20030124105A1
GENERAL INFORMATION:
APPLICANT: Miura, Junyong
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,455

FILING DATE: 2-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/368,704

FILING DATE: 4-JAN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/258,287

FILING DATE: 10-JUN-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/080,850

FILING DATE: 24-JUN-1993
ATTORNEY/AGENT INFORMATION:

NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086

REFERENCE/DOCKET NUMBER: 0609.3920003
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids

TYPE: amino acid
TOPOLOGY: both

US-08-459-455-36

Alignment Scores:
Pred. No.: 84.4 Length: 503
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 8 Gaps: 0

US-10-030-271-1 (1-909) x US-08-459-455-36 (1-503)

QY 215 AGCTCCAGCAGAGCTGTAGCCGCTG 189
DB 149 SerSerSerArgSerSerArgProLeu 157

RESULT 42

US-10-123-529-8

Sequence 8, Application US/10123529

Publication No. US20030105046A1

GENERAL INFORMATION:

APPLICANT: Yuan, Junyong
Friedlander, Robert M.

TITLE OF INVENTION: Interleukin Converting Enzyme (ICE)
and Central Nervous System Damage
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Keseler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/123,529
FILING DATE: 17-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/099,463A
FILING DATE: 18-Jun-1998
APPLICATION NUMBER: US 60/050,242
FILING DATE: 19-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609.4420001/JAG/LBB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-123-529-8
Alignment Scores:
Pred. No.: 84.4 Length: 503
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
Gaps: 0
DB: 14
US-10-030-271-1 (1-909) x US-10-123-529-8 (1-503)
QY 215 AGCTCGAGCAGAGCTTAGCCGCTG 189
Db 149 SerSerSerArgSerSerArgProLeu 157
RESULT 43
US-09-888-243-5
Sequence 5, Application US/09888243
Patent No. US20020136714A1
GENERAL INFORMATION:
APPLICANT: Horvitz, H. Robert
APPLICANT: Yuan, Junying
APPLICANT: Shaham, Shai
TITLE OF INVENTION: Relatedness of Human Interleukin-1beta
TITLE OF INVENTION: Convertase Gene to a C. Elegans Cell Death Gene, Inhibitory
TITLE OF INVENTION: Portions of these Genes and Uses Therefor
FILE REFERENCE: 01997/211003
CURRENT APPLICATION NUMBER: US/09/888,243
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: US 09/083,662
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: US 08/394,189
PRIOR FILING DATE: 1995-02-24

PRIOR APPLICATION NUMBER: US 08/282,211
PRIOR FILING DATE: 1994-07-11
PRIOR APPLICATION NUMBER: US 07/984,182
PRIOR FILING DATE: 1992-11-20
PRIOR APPLICATION NUMBER: US 07/897,788
PRIOR FILING DATE: 1992-06-12
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5
LENGTH: 505
TYPE: PRT
ORGANISM: Caenorhabditis briggsae
FEATURE:
NAME/KEY: VARIANT
LOCATION: 94, 95, 96, 120, 179, 318
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-888-243-5
Alignment Scores:
Pred. No.: 84.4 Length: 505
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
Gaps: 0
DB: 9
US-10-030-271-1 (1-909) x US-09-888-243-5 (1-505)
QY 215 AGCTCGAGCAGAGCTTAGCCGCTG 189
Db 151 SerSerSerArgSerSerArgProLeu 159
RESULT 44
US-10-425-114-62674
Sequence 62674, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO: 62674
LENGTH: 517
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3689-250-G2_F11.pcp
US-10-425-114-62674
Alignment Scores:
Pred. No.: 84 Length: 517
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
Gaps: 0
DB: 15
US-10-030-271-1 (1-909) x US-10-425-114-62674 (1-517)
QY 224 CGGCGCTCAAGCTTCAGCAGAGCTCT 198
Db 498 ArgArgSerSerSerArgSerSer 506
RESULT 45
US-10-425-114-42986

```
; Sequence 42986, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42986
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700470960_FLI.pep
US-10-425-114-42986

Alignment Scores:
Pred. No.: 82.5 Length: 564
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.00% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-1 (1-909) x US-10-425-114-42986 (1-564)

QY 224 CGGCGCTCCAGCTCCAGCAGGCTCT 198
Db 545 ArgArGserSerSerSerSerSerSer 553

RESULT 46
US-10-437-963-149484
; Sequence 149484, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149484
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(675)
; OTHER INFORMATION: unsure at all Xaa locations
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wainock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0

Alignment Scores:
Pred. No.: 79.5 Length: 675
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-149484 (1-675)

QY 505 AGACGGCGCGCGAGAGCGGCCCGCC 531
Db 133 ArgArGArGArGArGArGArGArGArG 141

RESULT 47
US-10-437-963-142715
; Sequence 142715, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142715
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(901)
; OTHER INFORMATION: unsure at all Xaa locations
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wainock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0

Alignment Scores:
Pred. No.: 74.9 Length: 901
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.97% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-1 (1-909) x US-10-437-963-142715 (1-901)

QY 663 GCGCGCGCGCGCGCGCGCGCGCGCG 689
Db 3 AlaProGlyAlaGlyAlaAlaAlaGly 11

RESULT 48
US-10-408-765A-1972
; Sequence 1972, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wainock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 1972
; LENGTH: 955
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1972

Alignment Scores:
Pred. No.: 74
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.00%
DB: 16
Gaps: 0

US-10-030-271-1 (1-909) x US-10-408-765A-1972 (1-955)

QY 224 CGCGCTCCAGCTCCAGCAGAGCTCT 198
DB 155 ArgArgSerSerSerArgSerSer 163

RESULT 49
US-10-276-774-2346
; Sequence 2346, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2346
; LENGTH: 962
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2346

Alignment Scores:
Pred. No.: 73.9
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 3.00%
DB: 15
Gaps: 0

US-10-030-271-1 (1-909) x US-10-276-774-2346 (1-962)

QY 224 CGCGCTCCAGCTCCAGCAGAGCTCT 198
DB 162 ArgArgSerSerSerArgSerSer 170

RESULT 50
US-10-282-122A-62686
; Sequence 62686, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari
; APPLICANT: Zybskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62686
; LENGTH: 2031
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (61)..(61)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (166)..(166)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (173)..(173)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (278)..(278)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (888)..(888)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (890)..(890)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (896)..(896)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (943)..(943)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (980)..(980)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1044)..(1044)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
```



```

NAME/KEY: MISC FEATURE
LOCATION: (1058) .. (1058)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1069) .. (1069)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1087) .. (1087)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1095) .. (1095)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1442) .. (1442)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1474) .. (1474)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1487) .. (1487)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1512) .. (1512)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1560) .. (1560)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1592) .. (1592)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1650) .. (1650)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1728) .. (1728)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1756) .. (1756)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1809) .. (1809)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1844) .. (1844)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1890) .. (1890)
OTHER INFORMATION: X=any amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1924) .. (1924)
OTHER INFORMATION: X=any amino acid
US-10-282-122A-62686

```

```

Alignment Scores:
Pred. No.: 63.3
Score: 9.00
Percent Similarity: 100.00%

```

```

Length: 2031
Matches: 9
Conservative: 0

```

```

Best Local Similarity: 100.00%
Query Match: 3.00%
DB: 15
US-10-030-271-1 (1-909) x US-10-282-122A-62686 (1-2031)
Qy 127 CCAGCTCGCAGTCGTCAGTTGCCCGC 101
Db 1246 ProAlaArgThrArgSerValAlaArg 1254

```

```

Search completed: March 22, 2005, 04:28:40
Job time: 263.761 secs

```

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 16:18:03 ; Search time 4091.48 Seconds
(without alignments)
10765.255 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctacgcggtcgcac.....tcagtgtgacgagctgcac 909

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

GenBank1: *
1: gb Da: *
2: gb Hcg: *
3: gb In: *
4: gb On: *
5: gb Ov: *
6: gb Pat: *
7: gb Ph: *
8: gb Pl: *
9: gb Pr: *
10: gb Ro: *
11: gb Sgs: *
12: gb Sy: *
13: gb Un: *
14: gb Vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	909	100.0	909	6	BD093311 Apoptosis
2	909	100.0	1883	6	BD093312 Apoptosis
3	909	100.0	1883	9	AK075328 Homo sapi
4	858	94.4	981	6	CQ719642 Sequence
5	858	94.4	981	6	AF443591 Homo sapi
6	858	94.4	1230	6	AX322754 Sequence
7	858	94.4	1230	6	AX431308 Sequence
8	756	83.2	1951	9	BC027930 Homo sapi
9	703	66.3	981	9	AF457575 Homo sapi
10	411	45.2	2012	9	AY125488 Homo sapi
11	383	42.1	2005	9	BC013372 Homo sapi
12	334	36.7	626	6	BD149917 Primer fo
13	334	36.7	626	6	AX869855 Sequence
14	322	35.4	1659	9	AK130203 Homo sapi
15	321	35.3	98663	9	AC010247 Homo sapi
16	281	30.9	1067	6	AR339362 Sequence
17	277	30.5	11084	9	AC006486 Homo sapi
18	252	27.7	303	6	AX431298 Sequence
19	187	20.6	1084	6	AX364851 Sequence

20	187	20.6	1106	6	AX364852 Sequence
21	177	19.5	603	6	AX136449 Sequence
22	177	19.5	603	6	BD123689 Secretary
23	144	15.8	16708	2	AC068283 Homo sapi
24	56	6.2	993	10	AF457576 Homo sapi
25	56	6.2	993	10	AF543541 Mus muscu
26	56	6.2	1540	2	BC037043 Mus muscu
27	56	6.2	10669	2	AC120393 Mus muscu
28	56	6.2	22699	2	AC079490 Mus muscu
29	23	2.5	1821	5	BC075935 Danio rer
30	23	2.5	125523	5	BC248385 Zebrafish
31	22	2.4	838	6	CQ411697 Sequence
32	22	2.4	141272	9	AP001976 Homo sapi
33	22	2.4	154937	2	AP002382 Homo sapi
34	21	2.3	323	6	CQ406469 Sequence
35	21	2.3	1553	8	AK060228 Oryza sat
36	21	2.3	66135	1	AP006840 Oryza sat
37	21	2.3	110000	1	AP006840 Oryza sat
38	21	2.3	110000	1	AP006840 Oryza sat
39	21	2.3	110000	1	AP006840 Oryza sat
40	21	2.3	110000	1	AP006840 Oryza sat
41	21	2.3	110000	1	AP006840 Oryza sat
42	21	2.3	110000	1	AP006840 Oryza sat
43	21	2.3	110000	1	AP006840 Oryza sat
44	21	2.3	110000	1	AP006840 Oryza sat
45	21	2.3	110000	1	AP006840 Oryza sat
46	21	2.3	110000	1	AP006840 Oryza sat
47	21	2.3	110000	1	AP006840 Oryza sat
48	21	2.3	118141	8	AP005190 Oryza sat
49	21	2.3	164263	8	AP005515 Oryza sat
50	21	2.3	250084	2	AC117545 Mus muscu
51	21	2.3	347786	1	BX640438 Bordetella
52	21	2.3	349146	1	BX640424 Bordetella
53	21	2.2	21	6	BD093317 Apoptosis
54	20	2.2	22	6	BD093320 Apoptosis
55	20	2.2	547	11	BV043023 S212P6545
56	20	2.2	614	8	AK072505 Oryza sat
57	20	2.2	945	10	RATPAB15X MUSHOX
58	20	2.2	1083	10	AK104164 Oryza sat
59	20	2.2	1148	8	AK060329 Oryza sat
60	20	2.2	1166	8	AK071358 Oryza sat
61	20	2.2	2117	8	AK071358 Oryza sat
62	20	2.2	3880	1	AVIHODPABYC X63650 A.vinelandi
63	20	2.2	6008	1	AVIHODPABYC X63650 A.vinelandi
64	20	2.2	10417	1	AE004546 Pseudomon
65	20	2.2	13914	1	AE004546 Pseudomon
66	20	2.2	90976	2	AC147595 Mus muscu
67	20	2.2	110000	1	AE017282_31 Continuation (32 o
68	20	2.2	110000	1	BX255276_11 Continuation (32 o
69	20	2.2	110000	8	CR382130_34 Continuation (35 o
70	20	2.2	110892	9	AP000633 Homo sapi
71	20	2.2	125217	8	AC097627 Homo sapi
72	20	2.2	144557	2	AP004631 Oryza sat
73	20	2.2	148947	8	AP004397 Oryza sat
74	20	2.2	156933	8	AC125411 Genomic s
75	20	2.2	151424	2	AC018367 Homo sapi
76	20	2.2	162241	8	AF485811 Oryza sat
77	20	2.2	162495	9	AC015647 Homo sapi
78	20	2.2	166042	9	AP006289 Homo sapi
79	20	2.2	173877	10	AC099624 Mus muscu
80	20	2.2	182619	2	AC127591 Mus muscu
81	20	2.2	189050	1	AL646069 Ralstonia
82	20	2.2	192020	2	AC102311 Mus muscu
83	20	2.2	193277	10	AC104324 Mus muscu
84	20	2.2	202212	2	CNS01RGF Human chr
85	20	2.2	222830	2	AC134899 Mus muscu
86	20	2.2	229035	2	AC117925 Rattus no
87	20	2.2	230318	2	AC123468 Rattus no
88	20	2.2	234247	2	AC130866 Rattus no
89	20	2.2	251186	2	AC120750 Rattus no
90	20	2.2	255557	2	AC098609 Rattus no
91	20	2.2	258285	2	AC095235 Rattus no
92	20	2.2	270654	2	AC096084 Rattus no

```

c 93 20 2.2 277000 1 SC0939109 AL939109 Streptomy
94 20 2.2 300029 1 AE016787 AE016787 pseudomon
95 20 2.2 301925 1 AP005046 AP005046 Streptomy
c 96 20 2.2 346259 1 BX640435 BX640435 Bordetella
c 97 20 2.2 346287 1 BX640450 BX640450 Bordetella
c 98 19 2.1 360 6 AX794353 AX794353 Sequence
99 19 2.1 360 6 AX803613 AX803613 Sequence
c 100 19 2.1 377 11 G48211 G48211 SHGC-60419

```

ALIGNMENTS

```

RESULT 1
BD093311
LOCUS Apoptosis Relating Factor. 909 bp DNA linear PAT 27-AUG-2002
DEFINITION BD093311
ACCESSION BD093311.1 GI:2263899
VERSION WO 0104300-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Oca,T., Isogai,T., Nishikawa,T., Kawai,Y., Miyoshi,S. and Sato,S.
TITLE Apoptosis Relating Factor
JOURNAL Patent: WO 0104300-A 1 18-JAN-2001;
HELIK RESEARCH INSTITUTE, TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA,
YURI KAWAI, SOUSUKE MIYOSHI, SUSUMU SATO

```

COMMENT

```

OS Homo sapiens (human)
PN WO 0104300-A/1
PD 18-JAN-2001
PF 06-JUL-2000 WO 2000JP004516
PR 08-JUL-1999 JP 99P 194179,18-OCT-1999 US 60/159586 PI
TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI,
PI SUSUMU SATO

```

```

PC C12N15/12, C07K14/47, C12N5/10, C12N1/21, C12N1/19, C12N1/15, C12P21/
02, C07K16/18, C12P21/08, G01N33/53, G01N33/577
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

FEATURES

```

source
1..909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

ORIGIN

```

Query Match 100.0%; Score 909; DB 6; Length 909;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGAGATGCTTGACTAC 60
DB 1 ATGGCGCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGATGCTTGACTAC 60
QY 61 TACGGAGTCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 120
DB 61 TACGGAGTCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 120
QY 121 GAGCTGAGCTCTGGGCTTCTCTGCTGATGAGGCTCTGGCGCGCGGAGGCTTAGCC 180
DB 121 GAGCTGAGCTCTGGGCTTCTCTGCTGATGAGGCTCTGGCGCGCGGAGGCTTAGCC 180
QY 181 CGGGCCCGCAGCGGCTTAGAGCTCTCTGAGAGTGAAGCGCCGCGGAGCTGCGGAG 240
DB 181 CGGGCCCGCAGCGGCTTAGAGCTCTCTGAGAGTGAAGCGCCGCGGAGCTGCGGAG 240
241 AGCAACTGCGGGCTGCTGGGGCAACTCTCTGCGCTGCTGGCGCGCCACGACTGCTCGC 300

```

```

DB 241 AGCAACTGCGGGCTGCTGGGGCAACTCTCTGCGCTGCTGGCGCGCCACGACTGCTCGC 300
QY 301 CACCTGCGCGCAGCGGCGCCCGCCAGTGTCTCCAGAACGCTATAGCTATGCGACCTCC 360
DB 301 CACCTGCGCGCAGCGGCGCCCGCCAGTGTCTCCAGAACGCTATAGCTATGCGACCTCC 360
QY 361 AGCTCTTCAAGAGACAGAGGCTAGCTGCGCTGCGCTGCGGAGTCAAGCACTTCTGCA 420
DB 361 AGCTCTTCAAGAGAGACAGAGGCTAGCTGCGCTGCGCTGCGGAGTCAAGCACTTCTGCA 420
QY 421 AATTCTCAGAGGCTCAGTGGAGAGACAGGCTCCCGCCCAACCAAGGGGCGCGGAGT 480
DB 421 AATTCTCAGAGGCTCAGTGGAGAGAGGCTCCCGCCCAACCAAGGGGCGCGGAGT 480
QY 481 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGGGCGCCAGCGGACCCAG 540
DB 481 CGGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGGGCGCCAGCGGACCCAG 540
QY 541 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGACTGTGACATCCGACTC 600
DB 541 CAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAAGCAAGTGACTGTGACATCCGACTC 600
QY 601 CGGGTTCAAGAGTACTGCGAGCATGGGCGACGCTTGAGAGGAGGCGGTCATCCGG 660
DB 601 CGGGTTCAAGAGTACTGCGAGCATGGGCGACGCTTGAGAGGAGGCGGTCATCCGG 660
QY 661 CGGCCCCAGGCGCTGCGCGCGCGCAGCTGACGCTTTTGGCGAGCCACCGCAGTGC 720
DB 661 CGGCCCCAGGCGCTGCGCGCGCGCAGCTGACGCTTTTGGCGAGCCACCGCAGTGC 720
QY 721 TCAAGGACCTGCGGCTCTGTGTGTTGTGACATCAAGTTCTCAAGCTCTCTATCTGAC 780
DB 721 TCAAGGACCTGCGGCTCTGTGTGTTGTGACATCAAGTTCTCAAGCTCTCTATCTGAC 780
QY 781 GCCTTCGCGGCGCAGTACTGAGTGGGCGCCTGCTGAGGCGCTGCGGCGGCTGCTG 840
DB 781 GCCTTCGCGGCGCAGTACTGAGTGGGCGCCTGCTGAGGCGCTGCGGCGGCTGCTG 840
QY 841 ACTGAGGCCCTGCGAGAGGCTGTGGGCGGAGGCTGTTGCTGCTGCTGATGTGGAT 900
DB 841 ACTGAGGCCCTGCGAGAGGCTGTGGGCGGAGGCTGTTGCTGCTGCTGATGTGGAT 900
QY 901 GAGGCTGAC 909
DB 901 GAGGCTGAC 909

```

RESULT 2

```

BD093312 1883 bp DNA linear PAT 27-AUG-2002
LOCUS Apoptosis Relating Factor.
DEFINITION BD093312
ACCESSION BD093312.1 GI:22638900
VERSION WO 0104300-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

```

REFERENCE

```

AUTHORS Oca,T., Isogai,T., Nishikawa,T., Kawai,Y., Miyoshi,S. and Sato,S.
TITLE Apoptosis Relating Factor
JOURNAL Patent: WO 0104300-A 2 18-JAN-2001;
HELIK RESEARCH INSTITUTE, TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA,
YURI KAWAI, SOUSUKE MIYOSHI, SUSUMU SATO

```

COMMENT

```

OS Homo sapiens (human)
PN WO 0104300-A/2
PD 18-JAN-2001
PF 06-JUL-2000 WO 2000JP004516
PR 08-JUL-1999 JP 99P 194179,18-OCT-1999 US 60/159586 PI
TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI,
PI SUSUMU SATO

```

```

PC
C12N15/12, C07K14/47, C12N5/10, C12N1/21, C12N1/19, C12N1/15, C12P21/
02,
PC C07K16/18, C12P21/08, G01N33/53, G01N33/577
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1.1883
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
Source
ORIGIN
Query Match 100.0%; Score 909; DB 6; Length 1883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTGCCTGACTAC 60
Db 124 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTGCCTGACTAC 183
Qy 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 120
Db 184 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 243
Qy 121 GAGCTGAGCTCTGCGCTTCTGCTGATGATGAGGCTCCTGGCGCGCGAGGCTTACCC 180
Db 244 GAGCTGAGCTCTGCGCTTCTGCTGATGATGAGGCTCCTGGCGCGCGAGGCTTACCC 303
Qy 181 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGTGAAGCCCGCGGCGAGTGGCGGAG 240
Db 304 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGTGAAGCCCGCGGCGAGTGGCGGAG 363
Qy 241 AGCAACTGCTGGCTGCTGGGCAACTCTGCGCGTCTGCGCCGACGACCTGCTGCG 300
Db 364 AGCAACTGCTGGCTGCTGGGCAACTCTGCGCGTCTGCGCCGACGACCTGCTGCG 423
Qy 301 CACCTGGCGGCGAGCGGCGCGGCGAGTGTCTCAGAGCGCTATGAGTGGACCTCC 360
Db 424 CACCTGGCGGCGAGCGGCGCGGCGAGTGTCTCAGAGCGCTATGAGTGGACCTCC 483
Qy 361 AGCTCTTCAAGAGAGCAGAGGATGCTGCGCTGCGCCTGCGCAGTCAAGCATTTGCA 420
Db 484 AGCTCTTCAAGAGAGCAGAGGATGCTGCGCTGCGCCTGCGCAGTCAAGCATTTGCA 543
Qy 421 AATTCTCAGCAGGGTCACTGGGAGACAGGCTTCCCCCAACCAAGCGGCGAGGAGT 480
Db 544 AATTCTCAGCAGGGTCACTGGGAGACAGGCTTCCCCCAACCAAGCGGCGAGGAGT 603
Qy 481 CGGGCCCGGCGCAGTGTGTGTCAGAGCGGCGGAGAGGGGCCCAAGCGGCGAGCC 540
Db 604 CGGGCCCGGCGCAGTGTGTGTCAGAGCGGCGGAGAGGGGCCCAAGCGGCGAGCC 663
Qy 541 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAGGCAAGTGAAGCTGATCCGAGTCC 600
Db 664 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAGGCAAGTGAAGCTGATCCGAGTCC 723
Qy 601 CGGGTTGAGCAGATCTGCGAGAGTGGGCGAGCTTTGAGAGCAGGCGGTGATCCCG 660
Db 724 CGGGTTGAGCAGATCTGCGAGAGTGGGCGAGCTTTGAGAGCAGGCGGTGATCCCG 783
Qy 661 CGGGCCCGGCGGCGGCGGCGAGTGAAGTGTGTTGGGCGAGGCCACCGAGTCTGG 720
Db 784 CGGGCCCGGCGGCGGCGGCGAGTGAAGTGTGTTGGGCGAGGCCACCGAGTCTGG 843
Qy 721 TCAAGGAGCTGAGCTGTGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGGAC 780
Db 844 TCAAGGAGCTGAGCTGTGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGGAC 903
Qy 781 GCTTTCTGGGGCGACTAGTGGGCGCTGCTGACAGGCGCTCGGGGCGTGTCTCG 840
Db 904 GCTTTCTGGGGCGACTAGTGGGCGCTGCTGACAGGCGCTCGGGGCGTGTCTCG 963

```

```

Qy 841 ACTGAGCCCTGCGAGAGGCTGTGGCCGGGAGGCTGTGCTGCTGCTGCTGAT 900
Db 964 ACTGAGCCCTGCGAGAGGCTGTGGCCGGGAGGCTGTGCTGCTGCTGCTGAT 1023
Qy 901 GAGGCTGAC 909
Db 1024 GAGGCTGAC 1032

RESULT 3
AK075328
LOCUS
DEFINITION
Homo sapiens cDNA PSEC004 f1s, clone NT2RM1000558, weakly similar
to Homo sapiens death effector domain-containing testicular
molecule mRNA.
ACCESSION
AK075328
VERSION
AK075328.1 GI:22761346
KEYWORDS
oligo capping; f1s (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Hayashi,K.,
Ishii,S., Saito,K., Yamamoto,J., Wakamatsu,A., Nagai,T.,
Nakamura,Y., Nagahari,K., Sugano,S. and Isegai,T.
HRI human cDNA sequencing project
Unpublished
2 (bases 1 to 1883)
Direct Submission
Isegai,T. and Yamamoto,J.
Submitted (20-MAR-2002) Takao Isegai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomic@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass
sequencing, clone selection and full insert sequencing; Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction; Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
COMMENT
Location/Qualifiers
1.1883
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM1000558"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RM1"
/notes="cloning vector: pUC19FL3-mRNA from uninduced NT2
neuronal precursor cells"
124..1104
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC11551.1"
/db_xref="GI:22761347"
/translation="MAISGTPAPCWEDECLDYGMVLSHRMRYVYGGDTTECELE
LAFIDBAPGAGGLARAGSLBLELBERGGGGSNLLGLQVLAVLHDLPLH
ARRRRRVSPEPSSYGTSSSKRTGEGCRRRRSSSSANSQQQWETGSPPTKQRS
RCRPSGCGARRRRGAPAPPOOSEPAPSPSEGGVTCIDIRLVAAVEYEHGPALEQVA
SRBPQALROLDFVGATVLRSDLDSSVVCIDIKFSLSYLDAMFGDYLGGALLQALR
GVFLTEALRAVGRBAVRLVSVDEADVEAGRRRLLMEEGGRRPTEAS"

FEATURES
Source
ORIGIN
Query Match 100.0%; Score 909; DB 9; Length 1883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTGCCTGACTAC 60
Db 124 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGATGAGTGCCTGACTAC 183
Qy 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 120

```

Db	184	TACGGGATGCTGTGCTTCACTGATGTTCAAGGTGGGGCGGCAACTGACCGATGTC	243
QY	121	GAGCTGGAGGCTCTCGGACCTTTCTGTGTGATGAGGCTCTGGGCGCGCGGAGGCTTTAGCC	180
Db	244	GAGCTGGAGGCTCTGGGCTTTCTGTGTGATGAGGCTCTGGGCGCGCGGAGGCTTTAGCC	303
QY	181	CGGGCCCGCAGCGGCTTAGAGCTCTGTGTGAGCTGAGCGCGCGGCGAGTGCGGCGAG	240
Db	304	CGGGCCCGCAGCGGCTTAGAGCTCTGTGTGAGCTGAGCGCGCGGCGAGTGCGGCGAG	363
QY	241	AGCAACTGCGGGCTGTGGGGAACTCTGCGCGTGTGTCGGCCGCAAGACTGTGCCG	300
Db	364	AGCAACTGCGGGCTGTGGGGAACTCTGCGCGTGTGTCGGCCGCAAGACTGTGCCG	423
QY	301	CACCTGGGGCGCAAGCGGGCGCGGCACTGTCTCCAGAAAGCTATAGCTATAGCACTTCC	360
Db	424	CACCTGGGGCGCAAGCGGGCGCGGCACTGTCTCCAGAAAGCTATAGCTATAGCACTTCC	483
QY	361	AGCTCTTCAAAAGAGGACAGAGGGTATGCTCGCCGTGCGCTGGGAGTCAAGAGTTCTGCA	420
Db	484	AGCTCTTCAAAAGAGGACAGAGGGTATGCTCGCCGTGCGCTGGGAGTCAAGAGTTCTGCA	543
QY	421	AATTCTCGAGAGGTCAGTGGGAGACAGAGCTTCCGCCCAACCAAGCGGACGGCGAGT	480
Db	544	AATTCTCGAGAGGTCAGTGGGAGACAGAGCTTCCGCCCAACCAAGCGGACGGCGAGT	603
QY	481	CGGGGGCGGGCCAGTGTGTGTGCCAGAGCGGGCGGAGAGGGGGCCCCAGCGGCAATCCAG	540
Db	604	CGGGGGCGGGCCAGTGTGTGTGCCAGAGCGGGCGGAGAGGGGGCCCCAGCGGCAATCCAG	663
QY	541	CAGCAGTCAAGACCCGCGACACTTCTCTGAAAGCAAGTGACTGTGACATCCGGCTC	600
Db	664	CAGCAGTCAAGACCCGCGACACTTCTCTGAAAGCAAGTGACTGTGACATCCGGCTC	723
QY	601	CGGGTTCAGACAGACTACTGTCGAGCATGAGCCAGCTTGTGAGCAAGGCGTGTGATCCGG	660
Db	724	CGGGTTCAGACAGACTACTGTCGAGCATGAGCCAGCTTGTGAGCAAGGCGTGTGATCCGG	783
QY	661	CGGGCCCGCAGGGCTGAGCCGCGGACGTGAGCGTGTGTGGGCAAGGCCACCGCAATGTGGCG	720
Db	784	CGGGCCCGCAGGGCTGAGCCGCGGACGTGAGCGTGTGTGGGCAAGGCCACCGCAATGTGGCG	843
QY	721	TCAAGGGGACTGGGCTGTGGTGTGTGACATCAAGTTCTAGAGCTCTCCATCTGGAGC	780
Db	844	TCAAGGGGACTGGGCTGTGGTGTGTGACATCAAGTTCTAGAGCTCTCCATCTGGAGC	903
QY	781	GCTTTCGGGGGCACTACCTGATGTCGCGCTGTGACAGCGCTTGCGGGGCGTGTCTCTG	840
Db	904	GCTTTCGGGGGCACTACCTGATGTCGCGCTGTGACAGCGCTTGCGGGGCGTGTCTCTG	963
QY	841	ACTGAGGCCCTTGCAGAGGCTGTGGGCGGGAGAGGCTGTTGCGCTGTGCTAGTGTGAT	900
Db	964	ACTGAGGCCCTTGCAGAGGCTGTGGGCGGGAGAGGCTGTTGCGCTGTGCTAGTGTGAT	1023
QY	901	GAGGCTGAC 909	
Db	1024	GAGGCTGAC 1032	

RESULT 4					
LOCUS	CQ719642				
DEFINITION	CQ719642	981 bp	DNA	linear	PAT 03-FEB-2004
ACCESSION	Sequence 5576 from Patent WO02068579.				
VERSION	CQ719642				
KEYWORDS	CQ719642.1 GI:42280499				
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
AUTHORS					

TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL	Patent: WO 02068579-A 5576 06-SEP-2002; PE Corporation (NY) (US)
FEATURES	Location/Qualifiers
source	1..981
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
ORIGIN	

Query Match	94.4%	Score	858	DB	6	Length	981
Best Local Similarity	99.9%	Pred. No.	0				
Matches	908	Conservative	0	Mismatches	1	Indels	0
				Gaps			0

OY	1	ATGGGCGCTATCCGGGTGCACCCCGGCGCCGTCGTCTGGAGAGAGAAATGAGTGGCTTGACATAC	60
Db	1	ATGGGCGCTATCCGGGTGCACCCCGGCGCCGTCGTCTGGAGAGAGAAATGAGTGGCTTGACATAC	60
OY	61	TACGGGAATGCTGTGCTGCCTTACCCGTATGTTTCCAGAGGTGTGGGGGGCAATGCACCGAGTCC	120
Db	61	TACGGGAATGCTGTGCTGCCTTACCCGTATGTTTCCAGAGGTGTGGGGGGCAATGCACCGAGTCC	120
OY	121	GAGCTGGAGACTCCTTGACCCTTTCTGTCTGATGAGAGTCTCTGCGCGCCGCGAGAGCTTAGCC	180
Db	121	GAGCTGGAGACTCCTTGACCCTTTCTGTCTGATGAGAGTCTCTGCGCGCCGCGAGAGCTTAGGC	180
OY	181	CGGGCCGCGAGCGGACTTAGAGCTCTGTCTGAGCTTGAGACGCGCGCGGCAATGTGGCGGAG	240
Db	181	CGGGCCGCGAGCGGACTTAGAGCTCTGTCTGAGAGCTGAGACGCGCGCGGCAATGTGGCGGAG	240
OY	241	AGCAACCTGGGGCTGTGGGGGCAATCTCGGCGCGTGGCTGGGCGCGGCAAGACTGTGCGCGG	300
Db	241	AGCAACCTGGGGCTGTGGGGGCAATCTCGGCGCGTGGCTGGGCGCGGCAAGACTGTGCGCGG	300
OY	301	CACCTGGCGCGCAAGCGCGCGCGCGCAAGTGTCTCCAGAACGCTATAGCTATGAGCACTTCC	360
Db	301	CACCTGGCGCGCAAGCGCGCGCGCGCAAGTGTCTCCAGAACGCTATAGCTATGAGCACTTCC	360
OY	361	AGCTTTTCAAAAGAGAGACAGAGGATAGCTGTGCGGTGCGCGTCCGCGAGTCAAGCATTTTGTGA	420
Db	361	AGCTTTTCAAAAGAGAGACAGAGGATAGCTGTGCGGTGCGCGTCCGCGAGTCAAGCATTTTGTGA	420
OY	421	AATTCTGACAGGGGTGATGGGAGACAGAGTCCCCCCCAACCAAGCGGACAGGCGGAGT	480
Db	421	AATTCTGACAGGGGTGATGGGAGACAGAGTCCCCCCCAACCAAGCGGACAGGCGGAGT	480
OY	481	CGGGCCGGGCCCATGATGTGTGCCAGACGGCGCGAGAGAGGGGCCCAAGCGGCAACCCAG	540
Db	481	CGGGCCGGGCCCATGATGTGTGCCAGACGGCGCGAGAGAGGGGCCCAAGCGGCAACCCAG	540
OY	541	CAGCAGTCAAGACCCCGCCAGACTTCTCTGTAAAGCAAAAGTACCTGTGACATCCGGCTTC	600
Db	541	CAGCAGTCAAGACCCCGCCAGACTTCTCTGTAAAGCAAAAGTACCTGTGACATCCGGCTTC	600
OY	601	CGGGTTCAGACAGATCTACGAGCATGGGCGAGGCTTTGAGACAGAGGCGTGCATCCCG	660
Db	601	CGGGTTCAGACAGATCTACGAGCATGGGCGAGGCTTTGAGACAGAGGCGTGCATCCCG	660
OY	661	CGGGCCCAAGCGGTGGCGCGCGAGCTGGAAGTGTTTGGGACAGGCCAAGCAAGTGTCCGC	720
Db	661	CGGGCCCAAGCGGTGGCGCGCGAGCTGGAAGTGTTTGGGACAGGCCAAGCAAGTGTCCGC	720
OY	721	TCAAGGAGACTGTGGGTCTGTGTGTTGTGTACATCAAGTTCTCAGAGTCTTCCTATCTGGAC	780
Db	721	TCAAGGAGACTGTGGGTCTGTGTGTTGTGTACATCAAGTTCTCAGAGTCTTCCTATCTGGAC	780
OY	781	GCTTCTTGGGGGCACTTACCTGAGTGGCGGCTGTGCTGAGAGCCCTTGGGGGCGGTTCCTG	840
Db	781	GCTTCTTGGGGGCACTTACCTGAGTGGCGGCTGTGCTGAGAGCCCTTGGGGGCGGTTCCTG	840
OY	841	ACTGAGGCGCTTCGAGAGGCTGTGGGCGGGAGGCTGTTGCGCTGCTGCTCAAGTGTGAT	900

```

Db      841 ACTGAGCCCTGCGAGAGGCTGTGGCGCGAGGCTGTTCCTGCTGCTGATGTGGAT 900
Qy      901 GAGGCTGAC 909
Db      901 GAGGCTGAC 909

RESULT 5
AF443591
LOCUS   AF443591      981 bp mRNA linear PRI 25-FEB-2002
DEFINITION Homo sapiens death effector domain-containing DNA-binding protein 2
            (DEDD2) mRNA, complete cds.
ACCESSION AF443591
VERSION   AF443591.1 GI:17933266
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
TITLE     1. (bases 1 to 981)
          Roth W., Stenner-Liewen, F., Pawlowski, K., Godzik, A. and Reed, J. C.
          Identification and characterization of DEDD2, a death effector
          domain-containing protein
JOURNAL   J. Biol. Chem. 277 (9), 7501-7508 (2002)
MEDLINE   21850646
PUBMED    11741985
REFERENCE 2. (bases 1 to 981)
AUTHORS   Roth, W. and Reed, J. C.
TITLE     Direct Submission
          Submitted (05-NOV-2001) Burnham Institute, 10901 N. Torrey Pines
          Rd, La Jolla, CA 92037, USA
FEATURES
source
1. 981
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   1. .981
   /gene="DEDD2"
   1. .981
   /gene="DEDD2"
   /codon_start=1
   /product="death effector domain-containing DNA-binding
   protein 2"
   /protein_id="AA148220.1"
   /db_xref="GI:17933267"
   /translation="MALSGSTPAPCPWEDECLDYGMLSHRFEVVGQUTCECEL
   LAFLLDEAPGAGGLARSGLELLLEERRQCDSESNRLGQLRLVLRHDLPHL
   ARKRRLVSPERSYGTSSSKRTGSCRRRQSSANSQCGMETGSPPTKQRS
   RGRPSGARRRRRGAAPAPOQSEPARPSSEKVTCDIRLRVRAEYCEHGLEQVVA
   SRPQALAROLDVFGAATAVLRSRDLDSVDCIDFESLSYLDAMFGDYLGGALLQALR
   GVPLTEALREAVREAVRLVSYDEADYEARRRLMBEERGRRPTLAS"

ORIGIN
Query Match      94.4%; Score 858; DB 9; Length 981;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGGCGCTATTCGGGCTGACCCCGCGCGCTGTGGAGAGAGATGAGTCTGGACTAC 60
Db      1 ATGGCGCTATTCGGGCTGACCCCGCGCGCTGTGGAGAGAGATGAGTCTGGACTAC 60
Qy      61 TAGCGGATGCTGTGCTTCAACCGTATGTGAGAGTGTGGCGCGGCAACTGACCGAGTGC 120
Db      61 TAGCGGATGCTGTGCTTCAACCGTATGTGAGAGTGTGGCGCGGCAACTGACCGAGTGC 120
Qy      121 GAGCTGAGAGCTCTGGGCTTTCTCTGATGATGAGCTCTGGCGCGCGGAGAGCTTAGCC 180
Db      121 GAGCTGAGAGCTCTGGGCTTTCTCTGATGATGAGCTCTGGCGCGCGGAGAGCTTAGCC 180
Qy      181 CGGCGCCGACGCGGCTTAGAGAGCTCTGTGTGAGAGCTGAGCGCGCGGCACTGCGCGAG 240
Db      181 CGGCGCCGACGCGGCTTAGAGAGCTCTGTGTGAGAGCTGAGCGCGCGGCACTGCGCGAG 240

```

```

Qy      241 AGCAACCTGGGCTGTGCTGGGCAACTCTCGCGGTGTGGCCCGCCAGACTTGCCG 300
Db      241 AGCAACCTGGGCTGTGCTGGGCAACTCTCGCGGTGTGGCCCGCCAGACTTGCCG 300
Qy      301 CACCTGGCGCGAAGCGCGCCCGCCAGTGTCTTCAGACGCTATAGCACTTCC 360
Db      301 CACCTGGCGCGAAGCGCGCCCGCCAGTGTCTTCAGACGCTATAGCACTTCC 360
Qy      361 AGCTCTCAANAAGAGAGAGAGGTAGTGCACGTGCGCGGTGGAGTCAACAGTTTCA 420
Db      361 AGCTCTCAANAAGAGAGAGGTAGTGCACGTGCGCGGTGGAGTCAACAGTTTCA 420
Qy      421 AATTCTCAGAGAGGTCAAGTGAAGACAGAGCTTCCCGCCCAACCAAGCGGCGGAGT 480
Db      421 AATTCTCAGAGAGGTCAAGTGAAGACAGAGCTTCCCGCCCAACCAAGCGGCGGAGT 480
Qy      481 CGGCGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGCGCCAGCGCACCCAG 540
Db      481 CGGCGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGCGCCAGCGCACCCAG 540
Qy      541 CAGACGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAACGTGACATCCGCTC 600
Db      541 CAGACGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAACGTGACATCCGCTC 600
Qy      601 CGGCTTGAGACAGTACTGCGAGCATGAGCCAGCCTTGGAGAGAGGCGTGGCATCCCG 660
Db      601 CGGCTTGAGACAGTACTGCGAGCATGAGCCAGCCTTGGAGAGAGGCGTGGCATCCCG 660
Qy      661 CGGCCCCAGGCGCTGTGGCGCGAGCTGAGCGTGTGGGAGGCGACCGAGTGTGCGC 720
Db      661 CGGCCCCAGGCGCTGTGGCGCGAGCTGAGCGTGTGGGAGGCGACCGAGTGTGCGC 720
Qy      721 TCAGAGGACCTGGGCTTGTGTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 780
Db      721 TCAGAGGACCTGGGCTTGTGTTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 780
Qy      781 GCTTTGTGGGCGACACTGAGTGGCGGCGCTGTGAGGCGCTGCGGCGGCGTTCCTG 840
Db      781 GCTTTGTGGGCGACACTGAGTGGCGGCGCTGTGAGGCGCTGCGGCGGCGTTCCTG 840
Qy      841 ACTGAGCCCTGCGAGAGGCTGTGGCGGAGAGGCTTGTGCTGTGATGTGGAT 900
Db      841 ACTGAGCCCTGCGAGAGGCTGTGGCGGAGAGGCTTGTGCTGTGATGTGGAT 900
Qy      901 GAGGCTGAC 909
Db      901 GAGGCTGAC 909

RESULT 6
AX322754
LOCUS   AX322754      1230 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 4 from Patent WO0192527.
ACCESSION AX322754
VERSION   AX322754.1 GI:18093742
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
          Tang, Y. T., Azinza, Y., Yue, H., Burford, N., Ding, L., Elliott, V. S.,
          Patterson, C. and Baughn, M. R.
          Regulators of apoptosis
          Patent: WO 0192527-A 4 06-DEC-2001;
          Incyte Genomics, Inc. (US)
FEATURES
source
1. 1230
   /organism="Homo sapiens"
   /mol_type="unassigned DNA"
   /db_xref="taxon:9606"
   /note="Incyte ID No: 3102521CB1"

```

ORIGIN	
Query Match	94.4%; Score 858; DB 6; Length 1230;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 908; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	1 ATGGCCCTATCCGGGTGACACCCCGGCGCTGCTGGAGAGAGATGATGCTTGGACTAC 60
DB	127 ATGGCCCTATCCGGGTGACACCCCGGCGCTGCTGGAGAGAGATGATGCTTGGACTAC 186
QY	61 TACGGATGCTGTCGCTTCAACCGTATGTTGAGAGTGTGGCGGGAACGTGACCGAGTGC 120
DB	187 TACGGATGCTGTCGCTTCAACCGTATGTTGAGAGTGTGGCGGGAACGTGACCGAGTGC 246
QY	121 GAGCTGAGCTCTTGAGCTTTTGTGCTGATGAGGCTTCTGCGCGCCGCGAGGCTTAAGCC 180
DB	247 GAGCTGAGCTCTTGAGCTTTTGTGCTGATGAGGCTTCTGCGCGCCGCGAGGCTTAAGCC 306
QY	181 CGGGCCCGACGCGCTTAAGCTCTGCTGAGACTGAGAGCGCGCGGCGAGTGTCCGCGAG 240
DB	307 CGGGCCCGACGCGCTTAAGCTCTGCTGAGACTGAGAGCGCGCGGCGAGTGTCCGCGAG 366
QY	241 AGCAACTGCGGCTGCTGGGGCAACTCTGCGGCTGCGCCGCGCAAGACTGTGCGG 300
DB	367 AGCAACTGCGGCTGCTGGGGCAACTCTGCGGCTGCGCCGCGCAAGACTGTGCGG 426
QY	301 CACTGCGCGGCAAGCGCGCGCGCGCAAGTGTCTCAAGACGCTATAGCTATGCGACCTCC 360
DB	427 CACTGCGCGGCAAGCGCGCGCGCGCAAGTGTCTCAAGACGCTATAGCTATGCGACCTCC 486
QY	361 AGCTCTCAAGAGGACAGAGGGTAGCTGCGCTGCGCTGCGGAGTCAAGCACTTGTGCA 420
DB	487 AGCTCTCAAGAGGACAGAGGGTAGCTGCGCTGCGCTGCGGAGTCAAGCACTTGTGCA 546
QY	421 AATTCTCAGCAGGGTAGTGGAGACAGGCTCCCCCAACCAAGCGGCGCGAGAGT 480
DB	547 AATTCTCAGCAGGGTAGTGGAGACAGGCTCCCCCAACCAAGCGGCGCGAGAGT 606
QY	481 CGGGCCCGGCGCAGTGTGTGTGTCAGACGCGCGCGGAGAGGGGCCCAAGCCGACCCCA 540
DB	607 CGGGCCCGGCGCAGTGTGTGTGTCAGACGCGCGCGGAGAGGGGCCCAAGCCGACCCCA 666
QY	541 CAGCACTCAGAGCGCCCGCAGACCTTCTCTGAAGGCAAAAGTGACTGTGACATCCGGCTC 600
DB	667 CAGCACTCAGAGCGCCCGCAGACCTTCTCTGAAGGCAAAAGTGACTGTGACATCCGGCTC 726
QY	601 CGGGTTTGAGCAGAGTACTGCGAGCANTGGGCGAGCTTGGAGCAGGGCGTGGATCCCG 660
DB	727 CGGGTTTGAGCAGAGTACTGCGAGCANTGGGCGAGCTTGGAGCAGGGCGTGGATCCCG 786
QY	661 CGGCCCCAGGCGCTGCGCGCGGCGAGCTGAGCTGTGTTGGGCGAGGCCCAAGGATGCTGCGC 720
DB	787 CGGCCCCAGGCGCTGCGCGCGGCGAGCTGAGCTGTGTTGGGCGAGGCCCAAGGATGCTGCGC 846
QY	721 TCAAGGAGCCTGAGGCTGTGTGTTTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGC 780
DB	847 TCAAGGAGCCTGAGGCTGTGTGTTTGTGATCATCAAGTCTCAGAGCTCTCTATCTGAGC 906
QY	781 GCCTTCTGGGGGCACTAAGTGTGAGCGCCTGCTGCAAGGCCCTTGGCGGGCGTGTCTCG 840
DB	907 GCCTTCTGGGGGCACTAAGTGTGAGCGCCTGCTGCAAGGCCCTTGGCGGGCGTGTCTCG 966
QY	841 ACTGAGCGCTGCGAGAGGCTGTGTGGGCGGAGAGGCTGTGCTGCTGTCAAGTGTGAT 900
DB	967 ACTGAGCGCTGCGAGAGGCTGTGTGGGCGGAGAGGCTGTGCTGCTGTCAAGTGTGAT 1026
QY	901 GAGGCTGAC 909
DB	1027 GAGGCTGAC 1035

RESULT 7
AX431308

LOCUS	AX431308	1924 bp	DNA	linear	PAT 28-JUN-2002
DEFINITION	Sequence 17 from Patent WO0240680.				
ACCESSION	AX431308				
VERSION	AX431308.1	GI:21656177			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1				
JOURNAL	Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S. H., Reed, J. C., Roth, W. and Stenmer-Liwen, F.				
	Novel death domain proteins				
	Patent: WO 0240680-A 17 23-MAY-2002;				
FEATURES	BURNHAM INST (US)				
source	Location/Qualifiers				
	1..1924				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	91..>1044				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD37280.1"				
	/db_xref="GI:21656178"				
	/translation="MALSSSTPAPCWEDECDIYGMVLSHMFVVGQLETECELEL				
	IAFLIDEAPGAGGILAPARASGLELLLEERRQCEBSNRLGLQLRVLARLDLPHL				
	AKRRRPVSPERYSTSSSRTEGSSRRRQSSSNASQOGMETGSPPTKRRS				
	RGRPSGARRRRRGAAPAAFOQSEBPAPSEBGKVCIDILRVAREYCEHPALQOVA				
	SRPQALARQLDVFGQATVILSRDLDSVVCIDIKFSELSYLDAPFWDYLSGALLQALR				
	GVFLTEALREAVGREAVRLVSVDEADYEAGRRRLIMEEGS"				
ORIGIN					
Query Match	94.4%;	Score 858;	DB 6;	Length 1924;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 908;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
QY	1	ATGGCGCTATCCGGGTGACACCCCGGCGCTGCTGGAGAGAGATGATGCTTGGACTAC	60		
Db	91	ATGGCGCTATCCGGGTGACACCCCGGCGCTGCTGGAGAGAGATGATGCTTGGACTAC	150		
QY	61	TACGGATGCTGTCGCTTCAACCGTATGTTGAGAGTGTGGCGGGAACGTGACCGAGTGC	120		
Db	151	TACGGATGCTGTCGCTTCAACCGTATGTTGAGAGTGTGGCGGGAACGTGACCGAGTGC	210		
QY	121	GAGCTGAGCTCTTGAGCTTTTGTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAAGCC	180		
Db	211	GAGCTGAGCTCTTGAGCTTTTGTGCTGATGAGGCTCTGCGCGCGCGAGGCTTAAGCC	270		
QY	181	CGGGCCCGCAGCGGCTTAAGCTCTGCTGAGACTGAGACGCCCGGCGAGTGTCCGCGAG	240		
Db	271	CGGGCCCGCAGCGGCTTAAGCTCTGCTGAGACTGAGACGCCCGGCGAGTGTCCGCGAG	330		
QY	241	AGCAACTGCGGCTGCTGGGGCAACTCTGCGGCTGCTGCGCGCGCAAGACTGTGCGCG	300		
Db	331	AGCAACTGCGGCTGCTGGGGCAACTCTGCGGCTGCTGCGCGCGCAAGACTGTGCGCG	390		
QY	301	CACTTGGCGGCAAGCGGCGCGCGCGCAAGTGTCTCCAGAACGCTATAGCTTAAGCACTCC	360		
Db	391	CACTTGGCGGCAAGCGGCGCGCGCGCAAGTGTCTCCAGAACGCTATAGCTTAAGCACTCC	450		
QY	361	AGCTCTTCAAGAAGCAGAGGGTAGTCCGTCGCGCTGCGGCAAGTCAAGCACTTCTGCA	420		
Db	451	AGCTCTTCAAGAAGCAGAGGGTAGTCCGTCGCGCTGCGGCAAGTCAAGCACTTCTGCA	510		
QY	421	AATTCTCAGCAGGGTAGTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGGAGT	480		
Db	511	AATTCTCAGCAGGGTAGTGGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGGAGT	570		
QY	481	CGGGCCCGGCGCAAGTGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCCAAGCCGACCCCA	540		
Db	571	CGGGCCCGGCGCAAGTGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCCAAGCCGACCCCA	630		

QY 541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAGAGCAAGTGACCTGTGACATCCGGCTC 600
DB 631 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAGAGCAAGTGACCTGTGACATCCGGCTC 690
QY 601 CGGGTTGAGAGAGTCTGAGAGATGGGCGAGCTTGGAGCGAGGGGTGGCATCCCGG 660
DB 691 CGGGTTGAGAGAGTCTGAGAGATGGGCGAGCTTGGAGCGAGGGGTGGCATCCCGG 750
QY 661 CGGCGCCAGAGCGCTGAGCGCGGACAGCTGATGTTGGAGCGAGCCAGCGAGTCTGGC 720
DB 751 CGGCGCCAGAGCGCTGAGCGCGGACAGCTGATGTTGGAGCGAGCCAGCGAGTCTGGC 810
QY 721 TCAAGGAGACTGGGCTCTGTGTTGGATCAATCAAGTTCTCAGAGCTCTCTATCTGAGC 780
DB 811 TCAAGGAGACTGGGCTCTGTGTTGGATCAATCAAGTTCTCAGAGCTCTCTATCTGAGC 870
QY 781 GCCTTCTGGGGGAGTACTGATGAGGCGCCCTGCTGACAGGCGCCGCGGGGCTGTCTCTG 840
DB 871 GCCTTCTGGGGGAGTACTGATGAGGCGCCCTGCTGACAGGCGCCGCGGGGCTGTCTCTG 930
QY 841 ACTGAGGCGCTGAGAGAGCTGTGGGCGGAGAGCGTGTTCGCTGCTGTGATGTGAT 900
DB 931 ACTGAGGCGCTGAGAGAGCTGTGGGCGGAGAGCGTGTTCGCTGCTGTGATGTGAT 990
QY 901 GAGGCTGAC 909
DB 991 GAGGCTGAC 999

RESULT 8
LOCUS BC027930
DEFINITION Homo sapiens death effector domain containing 2, mRNA (cDNA clone MGC:35529 IMAGE:5199772), complete cds.
ACCESSION BC027930
VERSION BC027930.1 GI:20379818
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1951)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marzella, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappert, T.E., Brownstein, M.J., Udell, T.B., Toshyuk, S., Scheetz, T.E., Brownstein, M.J., Udell, T.B., Toshyuk, S., Carninci, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whilton, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, J.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.B., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL PUBMED 12477932
REFERENCE Strausberg, R.
AUTHORS Direct Submission
TITLE Submitted (08-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT MGC help desk

Email: cgabs-rt@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Masfello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantridop, S., Thomas, P.J., Touchman, J.W., Tourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAX Plate: 49 Row: b Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19923049.
Location/Qualifiers
1. 1951
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:35529 IMAGE:5199772"
/issue_type="Brain, adult, 6 pooled whole brains"
/clone_id="NIH_MGC_114"
/lab_host="DH10B"
/note="Vector: pCMV-Sport6"
1. 1951
/gene="DEBD2"
/note="synonym: FLAME-3"
/db_xref="LocusID:162989"
69. 1049
/gene="DEBD2"
/codon_start=1
/product="death effector domain-containing DNA binding protein 2"
/protein_id="AAH27930.1"
/db_xref="GI:20379819"
/db_xref="LocusID:162989"
/translation="MALSGSTPAPCWEDECDLYGMLSHRMEVYVGGOLTECEL LAFLLDAPGAGGLAARSGLELLEKRGQDSNLLVLGQLVLAHDLPLH ARRRRRVSPERSYGTSSSKRTSGCRRRSSSSANSQQGWTGSPPTKQRS RRRPGGARRRRGAPAPQOSEPARPSSEGVTCIRLRVAEYCEHGPALEQVA SRPQALARLDYVGQATVLRSDIGSVCDIKFSLSTYLDAPMGDYLGAALLQALR GVFLTELRARVARENRLVSDVDYEAAGRRLRLMBEGGRPTAS"

ORIGIN
Query Match 83.2%; Score 756; DB 9; Length 1951;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 906; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGCTGACCCCGCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
DB 69 ATGGCGCTATCCGGGCTGACCCCGCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 128
QY 61 TACGGAGTGTGTCGCTTCAACCGATGTTGAGGTGGTGGGCGGCAACTGACCGAGTGC 120
DB 129 TACGGAGTGTGTCGCTTCAACCGATGTTGAGGTGGTGGGCGGCAACTGACCGAGTGC 188
QY 121 GAGCTGAGGCTCTGAGGCTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAGCC 180
DB 189 GAGCTGAGGCTCTGAGGCTTCTGCTGATGAGGCTCTGAGCGCGCGGAGGCTTAGCC 248
QY 181 CGGCGCCGAGCGGCTTGAAGCTTCTGCTGAGGCTGAGAGCGCGCGGCGGAGTGGCGAG 240
DB 249 CGGCGCCGAGCGGCTTGAAGCTTCTGCTGAGGCTGAGAGCGCGCGGCGGAGTGGCGAG 308

Qy	241	AGCAACCTGGGGCTGCTGGGGGCAACTCTCTGGGCTGGGCGGCGCAACAATCTGTCG	300
Db	309	AGCAACTGCGGCTGCTGGGGCAACTCTCTGGGCTGGGCGGCGCAACAATCTGTCG	368
Qy	301	CACCTGGCGGCGCAAGCGGCGCGGCGCACTGTCTCTCAGAACGCTATAGCTATGCACTCC	360
Qy	361	AGCTTTTCAAAGAGGACAGAGGGTAACTGCTGCTGCGGTGGCGGATCAAGCATTTCTGCA	420
Db	429	AGCTTTTCAAAGAGGACAGAGGGTAACTGCTGCTGCGGTGGCGGATCAAGCATTTCTGCA	488
Qy	421	AATTCCTAGACAGGGTCACTGGGGAACAAGGCTCCCCCAACCAAGCGGCGGCGAGT	480
Qy	481	CGGGGCGGCGCCCAATGGTGTGTCCAGACGCGCGCGAGAGGGGCGCCAGCGCACCCAG	540
Db	549	CGGGGCGGCGCCCAATGGTGTGTGTCCAGACGCGCGCGAGAGGGGCGCCAGCGCACCCAG	608
Qy	541	CAGCAGTCAAGGCCCGCCAGACCTTCTCTGAAAGCAAAGTGACCTGTGACATCGGCTC	600
Db	609	CAGCAGTCAAGGCCCGCCAGACCTTCTCTGAAAGCAAAGTGACCTGTGACATCGGCTC	668
Qy	601	CGGTTTGGAGCAGAGTACTGTGGACATAGGGCGACCTTGGAGCAGAGGCGTGGCATCCCGG	660
Db	669	CGGTTTGGAGCAGAGTACTGTGGACATAGGGCGACCTTGGAGCAGAGGCGTGGCATCCCGG	728
Qy	661	CGGCGCCAGAGCGCTGGCGCGCGCAGCTGGACGTGTTTGGGCGAGCGCACCGCAGTCTGGC	720
Db	729	CGGCGCCAGAGCGCTGGCGCGCGCAGCTGGACGTGTTTGGGCGAGCGCACCGCAGTCTGGC	788
Qy	721	TCAAAGGACCTGGGCTGTGTGTTGTGACATCAAGATTCTGAGGCTCTCTATCTGGAC	780
Db	789	TCAAAGGACCTGGGCTCGGTGTTGTGACATCAAGATTCTGAGGCTCTCTATCTGGAC	848
Qy	781	GCCTTTCTGGGGGCACTACTTAATGATGGGCGCCCTGCTGACGAGCGCTGGGGGCTGTCTTG	840
Db	849	GCCTTTCTGGGGGCACTACTTAATGATGGGCGCCCTGCTGACGAGCGCTGGGGGCTGTCTTG	908
Qy	841	ACTAGAGCGCTCGCGAGAGCTGTGTGGGCGGGAGGCTGTTGCGCTGCTGGTCAATGGGAT	900
Db	909	ACTAGAGCGCTCGCGAGAGCTGTGTGGGCGGGAGGCTGTTGCGCTGCTGGTCAATGGGAT	968
Qy	901	GAGGCTTGAC 909	
Db	969	GAGGCTTGAC 977	
RESULT 9			
AF457575		981 bp	mRNA linear PRI 10-APR-2002
LOCUS			
DEFINITION			Homo sapiens death effector domain-containing protein FLAME-3
VERSION			(FLAME-3) mRNA, complete cds.
ACCESSION			AF457575
KEYWORDS			AF457575.1 GI:20126793
SOURCE			
ORGANISM			Homo sapiens (human)
REFERENCE			
AUTHORS			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 981) Zhan,Y., Hegde,R., Srinivasula,S.M., Fernandes-Alnemri,T. and Alnemri,E.S.
TITLE			Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the TRIF102 subunit of human transcription factor I11C
JOURNAL			Cell Death Differ. 9 (4), 439-447 (2002)
MEDLINE			21961615
PUBMED			11965497
REFERENCE			2 (bases 1 to 981)
AUTHORS			Alnemri,E.S.
TITLE			Direct Submission

FEATURES		Location/Qualifiers	
source			
JOURNAL	Submitted (11-DEC-2001) Microbiology and Immunology, Thomas Jefferson University, 233 S. 10th Street, Philadelphia, PA 19107, USA		
gene	1. .981 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="19" /map="19q13.2" 1. .981 /gene="FLAME-3" 1. .981 /gene="FLAME-3" /codon_start=1 /product="death effector domain-containing protein FLAME-3" /protein_id="AAI0835.1" /db_xref="GI:20126794" /translation="MAISGSTPAPCWEDEDTDYGMILNRMPEVVGQI.TEECEIE LAFILDPGAGGGLARSGLELLLEHERGQCDENSLRLGLQLRLAHLDLPH ARKRRPVSPPERSYGTSSSKRTESCRRRQSSSAN50QMETSPYKQRR RGPSPGARRRRGAPAPAPQOESPARPSEKATCDI RLRYRAVEYREHGLDEGV SRPQALAROLDVFGQATPAVLRSDLSVVDI KFESELSYDAEMGDIQSALLQAL GVFLTEALREAVREAVRLVSVDEADYEAGRRLLMEEGRPTAS"		
CDS			
Query Match	Similarity	Score	DB
Best Local	99.3%;	Pred. No. 6e-275;	Length 981;
Matches	903;	Conservative	0; Mismatches 6; Indels 0; Gaps 0;
QY	1 ATGGCGCTATCCGGGTGACCCCGGCCCCGCTGCTGGAGAGATGACTGCTGACATAC	60	
Db	1 ATGGCGCTATCCGGGTGACCCCGGCCCCGCTGCTGGAGAGATGACTGCTGACATAC	60	
QY	61 TAGCGGATGCTGTGCTTACCGTATGTTGAGGTGTGGGCGGCACTGACCGAGTGC	120	
Db	61 TAGCGGATGCTGTGCTTACCGTATGTTGAGGTGTGGGCGGCACTGACCGAGTGC	120	
QY	121 GAGCTGAGACTCTGGGCTTTCTGTGGATGAGGCTCTGGCGCGCGAGGCTTAACTC	180	
Db	121 GAGCTGAGACTCTGGGCTTTCTGTGGATGAGGCTCTGGCGCGCGAGGCTTAACTC	180	
QY	181 CGGGCCGCGACGGGCTTGAAGCTCTCTGAGTGTGAGCGCCCGGGGACATGTCACGAG	240	
Db	181 CGGGCCGCGACGGGCTTGAAGCTCTCTGAGTGTGAGCGCCCGGGGACATGTCACGAG	240	
QY	241 AGCAACCTGGGGCTGTGGGGCACTCTGGGCGTGGGCGCGCGACGACCTGTGCGG	300	
Db	241 AGCAACCTGGGGCTGTGGGGCACTCTGGGCGTGGGCGCGCGACGACCTGTGCGG	300	
QY	301 CACCTGGGCGCAAGCGGCGCGGCGAGTGTCTCCAGAACGTTAAGTATGCGACTCC	360	
Db	301 CACCTGGGCGCAAGCGGCGCGGCGAGTGTCTCCAGAACGTTAAGTATGCGACTCC	360	
QY	361 AGCTTTCAAGAGAGACAGAGGTAGTCCGCTGCGCGCTCGGCACTCAAGACTTTGCA	420	
Db	361 AGCTTTCAAGAGAGACAGAGGTAGTCCGCTGCGCGCGCTCGGCACTCAAGACTTTGCA	420	
QY	421 AATTCTCAGCAGGGTCAATGGGAGACAGGCTTCCCCCAACCAAGCGGCAAGGGGAGT	480	
Db	421 AATTCTCAGCAGGGTCAATGGGAGACAGGCTTCCCCCAACCAAGCGGCAAGGGGAGT	480	
QY	481 CGGGGCGGCGCAGTGTGTGTCCAGAGCGGCGGCGAGAGGGGCCCAAGCCGACCCAG	540	
Db	481 CGGGGCGGCGCAGTGTGTGTCCAGAGCGGCGGCGAGAGGGGCCCAAGCCGACCCAG	540	
QY	541 CAGCAGTCAGAGCCCGCAGACTCTTCTTGAAGGCAATGTACTGTGACATCCGCTTC	600	
Db	541 CAGCAGTCAGAGCCCGCAGACTCTTCTTGAAGGCAATGTACTGTGACATCCGCTTC	600	
QY	601 CGGGTTGAGCAGACTCTGAGCATGGGCGACCTTGGAGCAGAGGCGTGGCATCCGG	660	
Db	601 CGGGTTGAGCAGACTCTGAGCATGGGCGACCTTGGAGCAGAGGCGTGGCATCCGG	660	

Db 601 CGGTTTCAGCAGATACCGGAGCATGGCCAGCTTTGAGACAGGAGGCGTGGCATCCGG 660
Qy 661 CGGCCCCAGAGCGCTGGCCGAGCTGATGTTGGGACAGGCCACCGCATGCTGGC 720
Db 661 CGGCCCCAGAGCGCTGGCCGAGCTGATGTTGGGACAGGCCACCGCATGCTGGC 720
Qy 721 TCAGAGGACCTGGGCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTTCTTATCTGAC 780
Db 721 TCAGAGGACCTGGGCTGTGTGTGTGTGACATCAAGTTCTCAGAGCTTCTTATCTGAC 780
Qy 781 GCCTTGGGGGACATCTAGATGGCCGCTGTGACAGGCCCTGGGGGGGTGTCTTG 840
Db 781 GCCTTGGGGGACATCTAGATGGCCGCTGTGACAGGCCCTGGGGGGGTGTCTTG 840
Qy 841 ACTAGAGCCCTGCGAGAGGCTGTGGGCGGAGAGCTGTTCGCTGTGATGATGAT 900
Db 841 ACTAGAGCCCTGCGAGAGGCTGTGGGCGGAGAGCTGTTCGCTGTGATGATGAT 900
Qy 901 GAGGCTGAC 909
Db 901 GAGGCTGAC 909
RESULT 10
AY125488 2012 bp mRNA linear PRI 18-SEP-2002
LOCUS Homo sapiens death effector domain-containing DNA-binding 2
DEFINITION protein (DEBD2) mRNA, complete cds.
ACCESSION AY125488
VERSION AY125488.1 GI:22475163
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2012)
AUTHORS Lee,J.C., Schickling,O., Stegh,A.H., Oshime,R.G., Dinsdale,D.,
Cohen,G.M. and Peter,M.E.
DEBD regulates degradation of intermediate filaments during
apoptosis
J. Cell Biol. 158 (6), 1051-1066 (2002)
MEDLINE 2220288
PUBMED 1235123
REFERENCE 2 (bases 1 to 2012)
AUTHORS Peter,M.B.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2002) Ben May Cancer Institute, University of
Chicago, 924 E. 57th Street, R112, Chicago, IL 60637, USA
FEATURES
source location/Qualifiers
1..2012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2012
/gene="DEBD2"
84..1049
/gene="DEBD2"
/codon_start=1
/product="death effector domain-containing DNA-binding 2
protein"
/protein_id="AA095240.1"
/db_xref="GI:22475164"
/translation="MALSGSTPAPCWEDECLDYGMILSRMEFVVGQI.TEGELI
LAPILDERGAGGLARARSLLELEBERGOCDESNRLGLQILRLAHDILPHL
AKRRRPERISGTSSSKRTSGCRKRSSSSANSQSGSPTRQRSSGRPS
GGARRRGAAPQOQSEPARPSSEKVTDIRLRYAAYCEHPALQCVASRPQ
ALARQDIVGQATVLRERDLSGVVCDIKFSELSTYDAFMGDYLSGALQLRGVFLT
EALREAVGREAVRILVSVDEADYENAGRRRLIMEEGRRPTLAS"

ORIGIN
Query Match 45.2%; Score 411; DB 9; Length 2012;
Best Local Similarity 99.8%; Pred. No. 7.9e-184;
Matches 461; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 448 GCTTCCCCCAACCAAGCGGACGCGAGTGGGGCCGCCACAGTGTGTGCCAGA 507
Db 516 GCTTCCCCCAACCAAGCGGACGCGAGTGGGGCCGCCACAGTGTGTGCCAGA 575
Qy 508 CGGCGGCGAGAGGGGCCCGACGCCGACCCAGACAGTACAGAGCCGCCAGCTTCC 567
Db 576 CGGCGGCGAGAGGGGCCCGACGCCGACCCAGACAGTACAGAGCCGCCAGCTTCC 635
Qy 568 TCTGAAGGCAAGTACCTGTGATCCGGCTCCGGGTGAGAGAGTACTGGAGCAT 627
Db 636 TCTGAAGGCAAGTACCTGTGATCCGGCTCCGGGTGAGAGAGTACTGGAGCAT 695
Qy 628 GGGCCAGGCTTTGAGACAGAGGCGTGGACATCCGCGCGGCCCGACGCGTGGCAGCTG 687
Db 696 GGGCCAGGCTTTGAGACAGAGGCGTGGACATCCGCGCGGCCCGACGCGTGGCAGCTG 755
Qy 688 GACGTGTTGGGACAGGCCACCGACAGTGTGCTGCAAGGACCTGGGCTGTGTGTGT 747
Db 756 GACGTGTTGGGACAGGCCACCGACAGTGTGCTGCAAGGACCTGGGCTGTGTGTGT 815
Qy 748 GACATCAAGTCTGAGAGCTCTCTATCTGAGCGCTTCTGGGGGACTACTGATGAGC 807
Db 816 GACATCAAGTCTGAGAGCTCTCTATCTGAGCGCTTCTGGGGGACTACTGATGAGC 875
Qy 808 GCCCTGCTGACAGGCGCTGCGGGGCGTGTCTCTGACTGAGGCGCTGCGAGAGCTGTGGGC 867
Db 876 GCCCTGCTGACAGGCGCTGCGGGGCGTGTCTCTGACTGAGGCGCTGCGAGAGCTGTGGGC 935
Qy 868 CGGAGAGCTGTTCCCTGCTGCTGATGATGATGAGCTGAC 909
Db 936 CGGAGAGCTGTTCCCTGCTGCTGATGATGATGAGCTGAC 977
RESULT 11
BC013372
LOCUS Homo sapiens death effector domain containing 2, mRNA (cDNA clone
MGC:16414 IMAGE:3941907), complete cds.
ACCESSION BC013372
VERSION BC013372.2 GI:33872465
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2005)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Udell,T.B., Toshylyuk,I.S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Wuzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Kzywinski,M.I., Skalska,U., Smalins,D.E.,
Schermer,A., Schein,I.B., Jones,S.J. and Wertz,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL 12477932
PUBMED 12477932
REFERENCE 2 (bases 1 to 2005)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) National Institutes of Health, Mammalian

REMARK
COMMENT

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15426522.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadanesystemsbio.org
Anup Madan, Jessica Faney, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
Series: IRAL Plate: 25 Row: f Column: 8.
Location/Qualifiers

FEATURES
source

gene

CDS

1. .2005
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:16414 IMAGE:3941907"
/tissue_type="Lung, small cell carcinoma"
/clone_id="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .2005
/gene="DDEDD2"
/note="synonym: FLAME-3"
/db_xref="LocusID:162989"
75. .632
/gene="DDEDD2"
/codon_start=1
/product="DDEDD2 protein"
/protein_id="AAH13372.2"
/db_xref="GI:33872466"
/db_xref="LocusID:162989"
/translation="MALGSTPAPCAWEDECDLVYGMASHRMFEVVGQUTTECELELAFLLDAPGAGAGLARGSLLELLEERGGODESNLLGLQLFLVLRHDLPLHARRRVPSPRYSYGTSSSSSKRTGSCRRRRSSANSQGSPPTRKRRRRGRS
GGARRRRGAPAPQOSEPQITFL"

ORIGIN

Query Match 42.1%; Score 383; DB 9; Length 2005;
Best Local Similarity 99.8%; Pred. No. 1.6e-170;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
DB 75 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 134
QY 61 TACGGAGATGCTGCTTACCGATGTTGAGAGTGGAGCGGCGCAACTGACCGAGTGC 120
DB 135 TACGGAGATGCTGCTTACCGATGTTGAGAGTGGAGCGGCGCAACTGACCGAGTGC 194
QY 121 GAGCTGAGCTCTGGCTTCTGCTGATGAGGCTCTGGAGCGCGCGAGAGCTTACC 180
DB 195 GAGCTGAGCTCTGGCTTCTGCTGATGAGGCTCTGGAGCGCGCGAGAGCTTACC 254
QY 181 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGCGCAGTGGCGAG 240
DB 255 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGCGCAGTGGCGAG 314
QY 241 AGCAACTGCGGCTGCTGGGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 315 AGCAACTGCGGCTGCTGGGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 374
QY 301 CACCTGGGCGCAGCGGCGCGGCGCAGTGTCTCAGAGGCTTAGCTATAGGACCTCC 360
DB 375 CACCTGGGCGCAGCGGCGCGGCGCAGTGTCTCAGAGGCTTAGCTATAGGACCTCC 434

QY 361 AGCTTTCAAGAGACAGAGGGTAGTGGCCGCTGGCGAGTCAACAGTTCTGCA 420
DB 435 AGCTTTCAAGAGACAGAGGGTAGTGGCCGCTGGCGAGTCAACAGTTCTGCA 494
QY 421 AATTCTCAGCAGG 434
DB 495 AATTCTCAGCAGG 508

RESULT 12

BD149917

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

OS

PN

PD

PF

PI

PI

PI

PI

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

PC

QY 301 CACCTGGCGCAGACGGCGCGCCAGTGTCTCCAGAACGCTATAGCTATGGCACTTCC 360
Db 388 CACCTGGCGCGCCAGACGGCGCGCCAGTGTCTCCAGAACGCTATAGCTATGGCACTTCC 447
QY 361 AGCTCTTCAAGAGAGACAGAGGCTA 385
Db 448 AGCTCTTCAAGAGAGACAGAGGCTA 472

RESULT 13
AX869855 626 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 4760 from Patent EP1074617.
DEFINITION AX869855
ACCESSION AX869855
VERSION AX869855.1 GI:40024718
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Ota, T., Isegai, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 4760 07-FEB-2001;
Research Association for Biotechnology (JP)
Location/Qualifiers
1..626
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 36.7%; Score 334; DB 6; Length 626;
Best Local Similarity 99.7%; Pred. No. 3.3e-147;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 60
Db 88 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 147
QY 61 TACGGGATGCTGCTGCTTACCGTATGTTTCAGAGTGTGGCGGCGCACTGACCGAGTGC 120
Db 148 TACGGGATGCTGCTGCTTACCGTATGTTTCAGAGTGTGGCGGCGCACTGACCGAGTGC 207
QY 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGTCTCTGGCGCGCGGAGGCTTAGCC 180
Db 208 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGTCTCTGGCGCGCGGAGGCTTAGCC 267
QY 181 CGGGCGCGCAGCGGCTAGAGTCTGCTGAGAGCTGAGCGCGCGGCGCAGTGGCGGAG 240
Db 268 CGGGCGCGCAGCGGCTAGAGTCTGCTGAGAGCTGAGCGCGCGGCGCAGTGGCGGAG 327
QY 241 AGCAACTGCGGCTGCTGCTGGGCACTCTGCGCGCTGCTGCGCGCGCACTGCTGCGG 300
Db 328 AGCAACTGCGGCTGCTGCTGGGCACTCTGCGCGCTGCTGCGCGCGCACTGCTGCGG 387
QY 301 CACCTGGCGCGCAGCGCGCGCGCACTGCTGCTGCAAGACGCTATAGCTATGGCACTTCC 360
Db 388 CACCTGGCGCGCAGCGCGCGCGCACTGCTGCTGCAAGACGCTATAGCTATGGCACTTCC 447
QY 361 AGCTCTTCAAGAGAGACAGAGGCTA 385
Db 448 AGCTCTTCAAGAGAGACAGAGGCTA 472

RESULT 14
AK130203 1659 bp mRNA linear PRI 10-SEP-2003
LOCUS Homo sapiens cDNA FLJ26693 f18, clone MFG07953.
DEFINITION AK130203
ACCESSION AK130203
VERSION AK130203.1 GI:34526957

KEYWORDS oligo capping; f18 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Taahito, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H.,
Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T.,
Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A.,
Kawakami, B., Nagai, K., Isegai, T. and Sugano, S.
NEDO human cDNA sequencing project
Unpublished
TITLE NEDO human cDNA sequencing project
JOURNAL
REFERENCE 2 (bases 1 to 1659)
AUTHORS Sugano, S. and Suzuki, Y.
TITLE Direct Submission
JOURNAL
Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: f18@naa.ms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
Location/Qualifiers
1..1659
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MFG07953"
/cell_type="macrophage"
/clone_id="MPG"
/note="Cloning vector: pME18SFL3-differentiated by
oxidant"

ORIGIN
Query Match 35.4%; Score 322; DB 9; Length 1659;
Best Local Similarity 100.0%; Pred. No. 1.5e-141;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 588 TGACATCCGGTCCGGGTCGAGAGAGTACTGAGAGATGGGCGAGCTTGGAGCAGAG 647
Db 427 TGACATCCGGTCCGGGTCGAGAGAGTACTGAGAGATGGGCGAGCTTGGAGCAGAG 486
QY 648 CGTGGCATCCCGGCGGCGCCAGGCGCTGGCGCGCAGCTGAGCGTGTGGCGAGGCCAC 707
Db 487 CGTGGCATCCCGGCGGCGCCAGGCGCTGGCGCGCAGCTGAGCGTGTGGCGAGGCCAC 546
QY 708 CGCAGTGTGCTGCAAGGAGCCTGGGCTCTGTGGTTTGTGACATCAAGTTCTCAGAGCT 767
Db 547 CGCAGTGTGCTGCAAGGAGCCTGGGCTCTGTGGTTTGTGACATCAAGTTCTCAGAGCT 606
QY 768 CTCTATCTGAGACGCTTCTGGGCGGACTACTGAGTGGCGGCGCTGCGAGGCGCTGGG 827
Db 607 CTCTATCTGAGACGCTTCTGGGCGGACTACTGAGTGGCGGCGCTGCGAGGCGCTGGG 666
QY 828 GGGGCGTGTCTGACTGAGGCGCTGCGAGAGCTGTGGGCGGAGAGGCTGTCCCTGCT 887
Db 667 GGGGCGTGTCTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGAGGCTGTCCCTGCT 726
QY 888 GGTGAGTGTGATGAGGCTGAC 909
Db 727 GGTGAGTGTGATGAGGCTGAC 748

RESULT 15
AC010247 95663 bp DNA linear PRI 13-JUL-2002
LOCUS AC010247c
DEFINITION Homo sapiens chromosome 19 clone CTC-378H22, complete sequence.

ACCESSION AC010247 GI:21743752
VERSION AC010247.9
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 95663)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 95663)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 95663)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 95663)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 13, 2002 this sequence version replaced gi:21637454.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
NOTE: Small insert shatter library only 85964-86181. 226bp single
subclone 9479-9697.
FEATURES
source location/Qualifiers
1..95663
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-378H22"
join(9479..9697,85964..86181)
/note="NOTE: Small insert shatter library only
85964-86181. 226bp single subclone 9479-9697."
ORIGIN
Query Match 35.3%; Score 321; DB 9; Length 95663;
Best Local Similarity 100.0%; Pred. No. 2.5e-141;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 589 GACATCCGCGCTCCGGTTCAGACAGACTGCGAGATGGCCAGCTTGGAGCAGGCG 648
DB 90864 GACATCCGCGCTCCGGTTCAGACAGACTGCGAGATGGCCAGCTTGGAGCAGGCG 90925
QY 649 GTGGCATCCCGCGCGCCCGAGGCGCTGGCGCGCACTGACGCTTTGGGCGAGCCACC 708
DB 90924 GTGGCATCCCGCGCGCCCGAGGCGCTGGCGCGCACTGACGCTTTGGGCGAGCCACC 90865
QY 709 GCAATGCTGCGCTCAAGGAGCCTGGGCTGTGTTGTGATCAATGTTCTCAGAGCTC 768
DB 90864 GCAATGCTGCGCTCAAGGAGCCTGGGCTGTGTTGTGATCAATGTTCTCAGAGCTC 90805
QY 769 TCCATATCGACGCTTCTGGGGGCGACTGAGTGCGCGCTGTCAGAGCCCTGGG 828
DB 90804 TCCATATCGACGCTTCTGGGGGCGACTGAGTGCGCGCTGTCAGAGCCCTGGG 90745
QY 829 GGCCTGTTCTGACTGAGGCGCTTCCGAGAGGCTGTGGCGCGGAGGCTGTTGCTGCTG 888
DB 90744 GGCCTGTTCTGACTGAGGCGCTTCCGAGAGGCTGTGGCGCGGAGGCTGTTGCTGCTG 90685
QY 889 GTCAGTGTGATGAGGCTGAC 909
|||||

DB 90684 GTCAGTGTGATGAGGCTGAC 90664
RESULT 16
LOCUS AR339362 1067 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 853 from patent US 6569662.
ACCESSION AR339362
VERSION AR339362.1 GI:33726219
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1067)
AUTHORS Tang, Y.T., Zhou, P. and Dmanac, R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 853 27-MAY-2003;
FEATURES
source location/Qualifiers
1..1067
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 30.9%; Score 281; DB 6; Length 1067;
Best Local Similarity 99.3%; Pred. No. 4.5e-122;
Matches 431; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGCGTGCACCCCGCCCGCTGCTGGAGAGAGATGCTGACTAC 60
DB 151 ATGGCGCTATCCGCGTGCACCCCGCCCGCTGCTGGAGAGAGATGCTGACTAC 210
QY 61 TACGGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 120
DB 211 TACGGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGCAACTGACCGAGTGC 270
QY 121 GAGCTGAGGCTCTGAGCTTCTGCTGATGAGAGTCTGCGCGCGCGGAGGCTTATGCC 180
DB 271 GAGCTGAGGCTCTGAGCTTCTGCTGATGAGAGTCTGCGCGCGCGGAGGCTTATGCC 330
QY 181 CGGGCCCGCAGCGGCGCTAGAGCTTCTGCTGAGCTGAGCGCGCGGCGAGTGCAGG 240
DB 331 CGGGCCCGCAGCGGCGCTAAGGCTCTGCTGAGAGTGCAGCGCGCGGCGAGTGCAGG 390
QY 241 AGCAACTGCGGCTGCTGGGCGCACTCTCGGCTGTGGCCCGGCGGCACTGCTGCCG 300
DB 391 AGCAACTGCGGCTGCTGGGCGCACTCTCGGCTGTGGCCCGGCGGCACTGCTGCCG 450
QY 301 CACCTGGCGGCAAGGCGCGCGGCGAGTCTCCAGAGCTATAGCTATGAGCACTCC 360
DB 451 CACCTGGCGGCAAGGCGCGCGGCGAGTCTCCAGAGCTATAGCTATGAGCACTCC 510
QY 361 AGCTCTCAAGAGAGCAGAGGAGTGTGCGCTGCGCGCGGCGAGTCAAGCTTCTGCA 420
DB 511 AGCTCTCAAGAGAGCAGAGGAGTGTGCGCTGCGCGCGGCGAGTCAAGCTTCTGCA 570
QY 421 AATTTCAGCAGG 434
DB 571 AATTTCAGCAGG 584
RESULT 17
LOCUS AC006486/c 111084 bp DNA linear PRI 02-FEB-1999
DEFINITION Homo sapiens chromosome 19, BAC CIT-B-147B23 (BC85722), complete
sequence.
ACCESSION AC006486
VERSION AC006486.1 GI:4210498
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 111084)

AUTHORS Lamerdin, J.E., McCreedy, P.M., Skowronski, E., Viswanathan, V., Burkhardt-Schultz, K., Gordon, L., Dias, J., Ramirez, M., Stiliwagen, S., Phan, H., Velasco, N., Do, L., Regala, W., Terry, A., Gaines, J., Dangnam, L., Erler, A., Christensen, M., Georgescu, A., Avila, J., Liu, S., Altir, C., Andreise, T., Trankheim, M., Amico-Keller, G., Cofield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G., Krommiller, B., Arellano, A., Sanders, C., Ow, D., Nolan, M., Truong, S., Kobayashi, A., Olsen, A.S., and Carrano, A.V.

TITLE Sequence analysis of a 3.2 Mb region in 19q13.2 between CYP21 and D19S178

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 111084)

AUTHORS Lamerdin, J.E.

TITLE Direct Submission

JOURNAL Submitted (02-FEB-1999) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA

COMMENT Map and sequence oriented from q centromere to telomere. BAC 85722 (CIT-B-147B23) is separated from BAC 48537 (CIT-B-50E14) to the right by a gap of approximately 20 Kb. There is currently no neighboring clone centromeric of BAC 85722 in the sequencing queue. Additional chromosome 19 map and sequence information are available from: <http://www.bic.llnl.gov/genome/genome.html>.

FEATURES

source

```

1..111084
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2 between CYP21 and D19S178"
/clone="CIT-B-147B23"
/note="LNL clone name: BC85722"
403..547
/rpc_family="LINE2"
complement(707..805)
/rpc_family="MIR"
811..1115
/rpc_family="AluSx"
complement(1121..1244)
/rpc_family="MIR"
complement(1561..1695)
/note="predicted exon, program: graal2exons_human_1.3
frame: 0, quality: good, score: 52.000-(1656..1558) DDS
similarity to AA725300 a112b07.s1 Soares parathyroid tumor
NbhPA Homo sapiens cDNA clone 1342549 3', (466..369); 98%
identity."
complement(1247..2549)
/rpc_family="AluSg"
2639..2944
/rpc_family="AluSx"
3015..3255
/rpc_family="AluB"
3257..3340
/rpc_family="(GGGA)n"
3611..3912
/rpc_family="AluSp"
3913..3934
/rpc_family="AT-rich"
3939..4228
/rpc_family="AluSx"
complement(4241..4311)
/rpc_family="Alu/FLAM"
complement(4322..4618)
/rpc_family="AluB"
complement(4623..4743)
/rpc_family="FLAM-C"
4832..5013
/rpc_family="L1MC4"
5131..6020
/rpc_family="L1MC4"
6028..6328
/rpc_family="AluSx"
6341..6639
/rpc_family="AluB"

```

```

misc_feature
complement(6648..6915)
/note="BLASTN similarity to X97703 (1..267); match: 0.98,
score: 1.1e-95; database searched: nt; H.sapiens DNA for
highly polymorphic microsatellite marker"
repeat_region
/rpc_family="(CA)n"
complement(6986..7107)
/note="predicted exon, program: graal2exons_human_1.3,
frame: 0, quality: excellent, score: 88.000"
7851..7908
/rpc_family="LINE2"
8169..8329
/rpc_family="MIR"
complement(8535..8862)
/note="predicted exon, program: graal2exons_human_1.3,
frame: 0, quality: excellent, score: 77.000"
complement(8543..8765)
/note="VDS similarity to AA511237 vhs6d04.r1 Soares mouse
mammary gland NbmWG Mus musculus cDNA clone 890983 5';
(61..279); 88% identity."
complement(8742..8900)
/note="VDS similarity to overlapping ESTs:
AA218681 zq96g08.r1 Stragene NT2 neuronal precursor
937230 Homo sapiens cDNA clone 649886 5'; (57..217); 98%
identity."
complement(9487..9546)
/note="VDS similarity to overlapping ESTs:
AA511237 vhs6d04.r1 Soares mouse mammary gland NbmWG Mus
musculus cDNA clone 890983 5'; (1..60); 95%
identity.-AA218681 zq96g08.r1 Stragene NT2 neuronal
precursor 937230 Homo sapiens cDNA clone 649886 5';
(1..56); 98% identity."
complement(10156..10230)
/note="predicted exon, program: graal2exons_human_1.3,
frame: 0, quality: good, score: 54.000"
10606..10877
/rpc_family="AluSx"
10879..11173
/rpc_family="AluSp"
11174..11471
/rpc_family="AluSx"
11481..11543
/rpc_family="MIR"
13069..13205
/rpc_family="AluB"
13206..13506
/rpc_family="AluSg"
13509..13691
/rpc_family="AluB"
complement(13734..13818)
/rpc_family="Alu/FLAM"
complement(13847..13926)
/rpc_family="LINE2"
complement(13941..14239)
/rpc_family="AluSc"
complement(14631..14958)
/rpc_family="MIR"
15082..15214
/rpc_family="AluSc"
15215..15514
/rpc_family="AluY"
15515..15690
/rpc_family="AluSc"
15708..16033
/rpc_family="AluSx"
16259..16271
/note="predicted exon, program: graal2exons_human_1.3,
frame: 1, quality: excellent, score: 94.000-(16907..17213)
DDS similarity to AA853296 NHRBcae0403f1 Normal Human
Tribecular Bone Cells Homo sapiens cDNA clone
NHRBcae0403. Score: 566 Identity: 297/303 (98%)."
complement(16928..17055)
/rpc_family="(GGA)n"

```


TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 2 24-JAN-2002;
Compugen Ltd. (IL)
FEATURES
SOURCE location/Qualifiers
1..1084
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 20.6%; Score 187; DB 6; Length 1084;
Best Local Similarity 100.0%; Pred. No. 1.9e-77;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 687 GAGCGTGTGGGAGGCGGACCGGAGTGTCTGCTCAAGGAGACTGGGCTGTGTGTTG 746
DB 764 GAGCGTGTGGGAGGCGGACCGGAGTGTCTGCTCAAGGAGACTGGGCTGTGTGTTG 823
QY 747 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGGCGACTAAGTGTG 806
DB 824 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGGCGACTAAGTGTG 883
QY 807 CGCCCTGTGAGGCGCCCTGGGGGGGTGTCTCTGACTAGGCGCCCTGGGAGAGGCTGTGG 866
DB 884 CGCCCTGTGAGGCGCCCTGGGGGGGTGTCTCTGACTAGGCGCCCTGGGAGAGGCTGTGG 943
QY 867 CCGGGAG 873
DB 944 CCGGGAG 950
RESULT 20
AX364852 1106 bp DNA linear PAT 15-FEB-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0206315.
ACCESSION AX364852
VERSION AX364852.1 GI:18696741
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Mintz,L., Freilich,S. and Bernstein,J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 3 24-JAN-2002;
Compugen Ltd. (IL)
FEATURES
SOURCE location/Qualifiers
1..1106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 20.6%; Score 187; DB 6; Length 1106;
Best Local Similarity 100.0%; Pred. No. 1.9e-77;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 687 GAGCGTGTGGGAGGCGGACCGGAGTGTCTGCTCAAGGAGACTGGGCTGTGTGTTG 746
DB 786 GAGCGTGTGGGAGGCGGACCGGAGTGTCTGCTCAAGGAGACTGGGCTGTGTGTTG 845
QY 747 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGGCGACTAAGTGTG 806
DB 846 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGGCGACTAAGTGTG 905
QY 807 CGCCCTGTGAGGCGCCCTGGGGGGGTGTCTCTGACTAGGCGCCCTGGGAGAGGCTGTGG 866
DB 906 CGCCCTGTGAGGCGCCCTGGGGGGGTGTCTCTGACTAGGCGCCCTGGGAGAGGCTGTGG 965
QY 867 CCGGGAG 873
DB 966 CCGGGAG 972

RESULT 21
AX136449 603 bp DNA linear PAT 30-MAY-2001
LOCUS
DEFINITION Sequence 371 from Patent EP1067182.
ACCESSION AX136449
VERSION AX136449.1 GI:14272853
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: EP 1067182-A 371 10-JUN-2001;
Helix Research Institute (JP)
FEATURES
SOURCE location/Qualifiers
1..603
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 19.5%; Score 177; DB 6; Length 603;
Best Local Similarity 100.0%; Pred. No. 1.2e-72;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGAGGAGAGATGAGTGCCTGACTAC 60
DB 123 ATGGCGCTATCCGGGTGACCCCGCCCGCTGAGGAGAGATGAGTGCCTGACTAC 182
QY 61 TACGGAGATCTGTGCTTCAACCGATATGTTGAGAGTGTGTGGCGGCAACTGACGAGTGC 120
DB 183 TACGGAGATCTGTGCTTCAACCGATATGTTGAGAGTGTGTGGCGGCAACTGACGAGTGC 242
QY 121 GAGCTGAGAGCTCTGGGCTTCTGCTGATGAGGCTCTGGGCGGCGGAGGCTTA 177
DB 243 GAGCTGAGAGCTCTGGGCTTCTGCTGATGAGGCTCTGGGCGGCGGAGGCTTA 299
RESULT 22
BD123689 603 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION Secretory protein or membrane protein.
ACCESSION BD123689
VERSION BD123689.1 GI:23218634
KEYWORDS JP 2002017376-A/198.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.
TITLE Secretory protein or membrane protein
JOURNAL Patent: JP 2002017376-A 198 22-JAN-2002;
HELIIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017376-A/198
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253173
PI TOSHITO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YORI KAWAI,TOMOYASU
PI SUGIYAMA,
PI KOJI HAYASHI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10, C12P21/02,C12Q1/68//C12P21/08,C12N15/00,C12N5/00 CC
PC Secretory protein or membrane protein
FH Key Location/Qualifiers

FT source 1.603
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1.603
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 19.5%; Score 177; DB 6; Length 603;
Best Local Similarity 100.0%; Pred. No. 1.2e-72;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTGACCCCGCCCGTCTGGAGAGAGATGCTGACTAC 60
Db 123 ATGGCGCTATCCGGGTGACCCCGCCCGTCTGGAGAGAGATGCTGACTAC 182
QY 61 TACGGATGCTGTGCTTACCGTATTTGAGGTGTGGCGGCACTGACCGAGTGC 120
Db 183 TACGGATGCTGTGCTTACCGTATTTGAGGTGTGGCGGCACTGACCGAGTGC 242
QY 121 GAGCTGAGCTCCTGAGCTTCTGCTGATGAGGCTCTGAGCGCCGCGAGGCTTA 177
Db 243 GAGCTGAGCTCCTGAGCTTCTGCTGATGAGGCTCTGAGCGCCGCGAGGCTTA 299
RESULT 23
AC068283/c 167108 bp DNA linear HTG 30-AUG-2000
LOCUS Homo sapiens chromosome 19 clone RP11-117E13, WORKING DRAFT
DEFINITION
SEQUENCE, 36 unordered pieces.
AC068283
AC068283.4 GI:9945176
HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Waterson,R.H.
1 (bases 1 to 167108)
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 167108)
Waterson,R.H.
Direct Submission
Submitted (30-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 30, 2000 this sequence version replaced gi:9558638.
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH031E13
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 146259 bases at least Q40
Consensus quality: 153006 bases at least Q30
Consensus quality: 156230 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 163958; sum-of-contigs
Quality coverage: 3.94 in Q20 bases; agarose-fp
Quality coverage: 4.25 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 1540: contig of 1540 bp in length
1541 gap of unknown length
1641 contig of 1143 bp in length
1641 2783: gap of unknown length
2784 2883: gap of unknown length
2884 4448: contig of 1565 bp in length
4449 4548: gap of unknown length
4549 6251: contig of 1703 bp in length
6252 6351: gap of unknown length
6352 9081: contig of 2730 bp in length
9082 9181: gap of unknown length
9182 10460: contig of 1279 bp in length
10461 10560: gap of unknown length
10561 12863: contig of 2303 bp in length
12864 12963: gap of unknown length
12964 15536: contig of 2573 bp in length
15537 15636: gap of unknown length
15637 18260: contig of 2624 bp in length
18261 18360: gap of unknown length
18361 20681: contig of 2321 bp in length
20682 20781: gap of unknown length
20782 24034: contig of 3253 bp in length
24035 24134: gap of unknown length
24135 27581: contig of 3447 bp in length
27582 27681: gap of unknown length
27682 30154: contig of 2473 bp in length
30155 30254: gap of unknown length
30255 33662: contig of 3408 bp in length
33663 33762: gap of unknown length
33763 35826: contig of 2064 bp in length
35827 35926: gap of unknown length
35927 39520: contig of 3594 bp in length
39521 39620: gap of unknown length
39621 42496: contig of 2876 bp in length
42497 42596: gap of unknown length
42597 45290: contig of 2694 bp in length
45291 45390: gap of unknown length
45391 50042: contig of 4652 bp in length
50043 50142: gap of unknown length
50143 54382: contig of 4240 bp in length
54383 54482: gap of unknown length
54483 58846: contig of 4364 bp in length
58847 58946: gap of unknown length
58947 63662: contig of 4716 bp in length
63663 63762: gap of unknown length
63763 69752: contig of 5990 bp in length
69753 69852: gap of unknown length
69853 75209: contig of 5357 bp in length
75210 75309: gap of unknown length
75309 80744: contig of 5435 bp in length
80745 80844: gap of unknown length
80845 87050: contig of 6206 bp in length
87051 87150: gap of unknown length
87151 92697: contig of 5547 bp in length
92698 92797: gap of unknown length
92798 96711: contig of 3914 bp in length
96712 96811: gap of unknown length
96812 103572: contig of 6761 bp in length
103573 103672: gap of unknown length
103673 111357: contig of 7685 bp in length
111358 111457: gap of unknown length
111458 119120: contig of 7663 bp in length
119121 119220: gap of unknown length
119221 126481: contig of 7261 bp in length
126482 126582: gap of unknown length
126582 138776: contig of 12195 bp in length
138777 138876: gap of unknown length
138877 154763: contig of 15887 bp in length
154764 154863: gap of unknown length

FEATURES	
source	
* 154864	166124: contig of 11261 bp in length
* 166125	166224: gap of unknown length
* 166225	167108: contig of 884 bp in length.
Location/Qualifiers	
1..167108	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	
/chromosome="19"	
/clone="RP11-317E13"	
1..1540	
misc_feature	/note="assembly_name:Contig23"
misc_feature	1641..2783
misc_feature	/note="assembly_name:Contig24"
misc_feature	2884..4448
misc_feature	/note="assembly_name:Contig25"
misc_feature	4549..6251
misc_feature	/note="assembly_name:Contig26"
misc_feature	6352..9081
misc_feature	/note="assembly_name:Contig27"
misc_feature	9182..10460
misc_feature	/note="assembly_name:Contig28"
misc_feature	10561..12863
misc_feature	/note="assembly_name:Contig29"
misc_feature	12964..15336
misc_feature	/note="assembly_name:Contig30"
misc_feature	15637..18260
misc_feature	/note="assembly_name:Contig31"
misc_feature	18361..20681
misc_feature	/note="assembly_name:Contig32"
misc_feature	20782..24034
misc_feature	/note="assembly_name:Contig33"
misc_feature	24135..27581
misc_feature	/note="assembly_name:Contig34"
misc_feature	27682..30154
misc_feature	/note="assembly_name:Contig35"
misc_feature	30255..33662
misc_feature	/note="assembly_name:Contig36"
misc_feature	33763..35826
misc_feature	/note="assembly_name:Contig37"
misc_feature	35927..39520
misc_feature	/note="assembly_name:Contig38"
misc_feature	39621..42496
misc_feature	/note="assembly_name:Contig39"
misc_feature	42597..45290
misc_feature	/note="assembly_name:Contig40"
misc_feature	45391..50042
misc_feature	/note="assembly_name:Contig41"
misc_feature	50143..54382
misc_feature	/note="assembly_name:Contig42
misc_feature	clone end:SP6
misc_feature	vector_side:left"
misc_feature	54483..58846
misc_feature	/note="assembly_name:Contig43"
misc_feature	58947..63662
misc_feature	/note="assembly_name:Contig44"
misc_feature	63763..69752
misc_feature	/note="assembly_name:Contig45"
misc_feature	69853..75209
misc_feature	/note="assembly_name:Contig46"
misc_feature	75310..80744
misc_feature	/note="assembly_name:Contig47"
misc_feature	80845..87050
misc_feature	/note="assembly_name:Contig48"
misc_feature	87151..92637
misc_feature	/note="assembly_name:Contig49"
misc_feature	92798..96711
misc_feature	/note="assembly_name:Contig50"
misc_feature	96812..103572
misc_feature	/note="assembly_name:Contig51"
misc_feature	103673..111357
misc_feature	/note="assembly_name:Contig52"
misc_feature	111458..119120

misc_feature	/note="assembly_name:Contig53"
misc_feature	119221..126481
misc_feature	/note="assembly_name:Contig54"
misc_feature	126582..138776
misc_feature	/note="assembly_name:Contig55"
misc_feature	138877..154763
misc_feature	/note="assembly_name:Contig56"
misc_feature	154864..166124
misc_feature	/note="assembly_name:Contig57"
Query Match	15.8%; Score 144; DB 2; Length 167108;
Best Local Similarity	100.0%; Pred. No. 2.6e-57; Indels 0; Gaps 0;
Matches 144; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	446 CAGGCTCCCCCAACCAAGCGGAGCGGCGGAGTCGGGCGGCGCCAGTGTGTGCCA 505
Db	23027 CAGGCTCCCCCAACCAAGCGGAGCGGCGGAGTCGGGCGGCGCCAGTGTGTGCCA 22966
QY	506 GACGCGCGCGGAGGCGGCGCCAGCGCCATCCCAAGCAGCAGTCAAGAGCCGCCAGACTT 565
Db	22967 GACGCGCGCGGAGGCGGCGCCAGCGCCATCCCAAGCAGCAGTCAAGAGCCGCCAGACTT 22900
QY	566 CCTCTGAAGCGCAAGTGACCGTGG 589
Db	22907 CCTCTGAAGCGCAAGTGACCTGTG 22884
RESULT 24	
LOCUS	AF457576 993 bp mRNA linear ROD 10-APR-2002
DEFINITION	Mus musculus death effector domain-containing protein FLAME-3
ACCESSION	AF457576
VERSION	AF457576
KEYWORDS	AF457576.1 GI:20126795
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 993)
REFERENCE	Zhan, Y., Hegde, R., Srinivasula, S.M., Fernandes-Alnemri, T. and Alnemri, E.S. Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the TP53 p102 subunit of human transcription factor p53
REFERENCE	Cell Death Differ. 9 (4), 439-447 (2002)
TITLE	Journal MEDLINE 21961615
TITLE	PubMed 11965497
TITLE	2 (bases 1 to 993)
TITLE	Alnemri, E.S.
TITLE	Direct Submission
TITLE	Submitted (11-DEC-2001) Microbiology and Immunology, Thomas Jefferson University, 233 S. 10th Street, Philadelphia, PA 19107, USA
FEATURES	
source	location/Qualifiers
source	1..993
source	/organism="Mus musculus"
source	/mol_type="mRNA"
source	/strain="C57BL/6J"
source	/db_xref="taxon:10090"
source	/clone="flip"
source	1..993
source	/gene="Flame-3"
source	1..993
source	/gene="Flame-3"
source	/codon_start=1
source	/product="death effector domain-containing protein FLAME-3"
source	/protein_id="AA010836.1"
source	/db_xref="GI:20126796"
source	/translation="MALSGSTPAAPSEBEDCDYGMLSLHRFVVGQLTCELEL LAFLLDAPAGPGLARSGLELLELRGGCDSSNRLFLSQLRLVLRDLPLHL ARKRRRYSERYSYGNPSSSRTTDSGCRQASSSSDSPQSQWDTGSPPTQRQR

ORIGIN

Query Match 6.2%; Score 56; DB 10; Length 993;
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

769 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 824
 781 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 836

SRGRSSGARRRRRAGLAASQOHQHOELGRSSSEKVTCDIHLRVRAECHEGPALE
 QGVASRRPQALRQLDVFCQATVRLVSRDGSVVCIIKPSSELSYLDAPMDYISGALL
 QALRGVFLTEALREAVGREAVRLVSVDEADYEAARRLLMEEGRRGTENS"

ORIGIN

Query Match 6.2%; Score 56; DB 10; Length 993;
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

769 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 824
 781 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 836

AF543541 993 bp mRNA linear ROD 02-OCT-2002
 AF543541
 AF543541
 AF543541 GI:23452816

DEFINITION Mus musculus death effector domain-containing DNA-binding protein 2 (Dedd2) mRNA, complete cds.

ACCESSION AF543541

VERSION AF543541

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 993)

REFERENCE Alciavar, A., Hu, S. and Yang, X. DEDD2 is a DED-containing protein targeted to the nucleolus

TITLE DEDD2 is a DED-containing protein targeted to the nucleolus

AUTHORS Alciavar, A., Hu, S. and Yang, X.

REFERENCE 2 (bases 1 to 993)

JOURNAL Direct Submission

AUTHORS Submitted (06-SEP-2002) Abramson Family Cancer Research Institute, Univ. of Pennsylvania, 421 Curie Blvd. 635 BRB1/III, Philadelphia, PA 19104-6160, USA

JOURNAL Location/Qualifiers

FEATURES

source

1. .993

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

1. .993

/gene="Dedd2"

/gene="Dedd2"

/codon_start=1

/product="death effector domain-containing DNA-binding protein 2"

/protein_id="AA033179.1"

/db_xref="GI:23452816"

/translation="MALSGSTPAWSWEDECLDYGMLSHRMEVVGQGLTECELEL
 LAFILDEAPAGGLARARGLLELERRGQCDSENLILSLQLRLVLRHLLPHL
 ARRRRVPSPERYSGNPSKRTESRRROASSSSPSOMDTPPKROR
 SRGRSSGARRRRRAGLAASQOHQHOELGRSSSEKVTCDIHLRVRAECHEGPALE
 QGVASRRPQALRQLDVFCQATVRLVSRDGSVVCIIKPSSELSYLDAPMDYISGALL
 QALRGVFLTEALREAVGREAVRLVSVDEADYEAARRLLMEEGRRGTENS"

ORIGIN

Query Match 6.2%; Score 56; DB 10; Length 993;
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

769 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 824
 781 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 836

SRGRSSGARRRRRAGLAASQOHQHOELGRSSSEKVTCDIHLRVRAECHEGPALE
 QGVASRRPQALRQLDVFCQATVRLVSRDGSVVCIIKPSSELSYLDAPMDYISGALL
 QALRGVFLTEALREAVGREAVRLVSVDEADYEAARRLLMEEGRRGTENS"

ORIGIN

Query Match 6.2%; Score 56; DB 10; Length 993;
 Best Local Similarity 100.0%; Pred. No. 3.1e-15;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

769 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 824
 781 TCCTATCTGACCGCTTCTGGGGCGACTGAGTGGCGCCCTGCTGCAGGCCCT 836

SRGRSSGARRRRRAGLAASQOHQHOELGRSSSEKVTCDIHLRVRAECHEGPALE
 QGVASRRPQALRQLDVFCQATVRLVSRDGSVVCIIKPSSELSYLDAPMDYISGALL
 QALRGVFLTEALREAVGREAVRLVSVDEADYEAARRLLMEEGRRGTENS"

VERSION BC037043.1 GI:22477602

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1540)

AUTHORS Strausberg, R., Feingold, E. A., Grouse, L. H., Derge, J. G., Klauener, R. D., Collins, S. S., Wagner, L., Shenmen, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Heich, F., Diatchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Uedlin, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullany, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Villalon, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smailus, D. E., Scherch, A., Schein, J. E., Jones, S. J. and Marra, M. A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1540)

Strausberg, R.

Direct Submission

Submitted (23-AUG-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-rc@mail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),

Galtherisburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhtar, N., Ayale, K., Beckstrom-Sternberg, S. M., Benjamin, B., Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N. L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaepi, R., Maduro, Q. L., Mastello, C., Maskeri, B., Mastrian, S. D., McCloskey, J. C., McDowell, J., Pearson, R., Stantripod, S., Thomas, P. J., Touchman, J. W., Tsurgoun, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E. D.

REMARK

COMMENT

1. .1540

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4526883"

/tissue_type="Eye, retina, mouse strain C57Bl/6"

/clone_lib="NIH_MGC_94"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

FEATURES

source

1. .1540

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:4526883"

/tissue_type="Eye, retina, mouse strain C57Bl/6"

/clone_lib="NIH_MGC_94"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

gene
1. 1540
/gene="Ded22"
/note="synonym: FLAME-3"
/db_xref="locusID:67379"
/db_xref="MGI:1914629"
281.754
/gene="Ded22"
/codon_start=1
/product="Ded22 protein"
/protein_id="AAH37043.1"
/db_xref="GI:22477603"
/db_xref="LocusID:67379"
/db_xref="MGI:1914629"
/translation="MGKQAKRGVTCRPTARAWQSDLSIDRLRYAEYCEHGPAL
EQGVASRRPOLARQLDFGATVLRSDISVCDIKFSELSYLAFMGDYLSGAL
LQALRGVFLTEALREAVGREAIVRLVSVDEADYAGRRRLILMEEGRGRTGAS"

ORIGIN

Query Match 6.2%; Score 56; DB 10; Length 1540;
Best Local Similarity 100.0%; Pred. No. 2.9e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

cy 769 TCCATCTGACGCGCTCTGGGGGACATCACTGAGCGCCCTGCTGCAGCCCT 824
Db 542 TCCATCTGACGCGCTCTGGGGGACATCACTGAGCGCCCTGCTGCAGCCCT 597

RESULT 27
AC120393/c 190669 bp DNA linear HTG 05-JUL-2004
LOCUS Mus musculus chromosome 7 clone RP24-312B12 map 7, *** SEQUENCING
DEFINITION IN PROGRESS ***, 19 unordered pieces.
AC120393
AC120393.4 GI:45736765
HTG; HTGS PHASE1; HTGS_FUJLTOP; HTGS_ACTIVEFIN.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus chromosome 7, clone RP24-312B12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 190669)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collimore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Ryback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strausen, N., Subramanian, A., Talama, J., Testfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmerman, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

COMMENT
TITLE
JOURNAL
REFERENCE
AUTHORS

Choepe, Y., Collymore, A., Cooke, P., Corum, B.,
Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L.,
Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D.,
Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N.,
Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T.,
Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R.,
Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C.,
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C.,
Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Stubbs, M., Talama, J., Testfaye, S., Theodore, J.,
Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R.,
Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L.,
Zimmer, A. and Zody, M.

Direct Submission
Submitted (05-JUL-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 25, 2004 this sequence version replaced gi:30017810.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WITR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
Project Information
Center project name: L25953
Center clone name: 312_B_12

NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
23038: contig of 23038 bp in length
23039 23138: gap of 100 bp
23139 25827: contig of 2689 bp in length
25827 25928: gap of 100 bp
25928 28740: contig of 2813 bp in length
28740 28841: gap of 100 bp
28841 45485: contig of 16645 bp in length
45485 45585: gap of 100 bp
45585 48996: contig of 3411 bp in length
48996 49095: gap of 100 bp
49095 53311: contig of 3215 bp in length
53311 52411: gap of 100 bp
52411 55723: contig of 3312 bp in length
55723 55823: gap of 100 bp
55823 69848: contig of 14025 bp in length
69848 69949: gap of 100 bp
69949 83553: contig of 13605 bp in length
83553 83653: gap of 100 bp
83653 96037: contig of 12384 bp in length
96037 96137: gap of 100 bp
96137 101113: contig of 4976 bp in length
101113 101213: gap of 100 bp
101213 103726: contig of 2513 bp in length
103726 103827: gap of 100 bp
103827 106119: contig of 2293 bp in length
106119 106219: gap of 100 bp
106219 116675: contig of 10456 bp in length
116675 116775: gap of 100 bp
116775 127768: contig of 10993 bp in length
127768 127869: gap of 100 bp
127869 135609: contig of 7741 bp in length

FEATURES
source
1. 190669
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/map="7"
/clone="RP24-312B12"
/clone_11b="RPCL-24 Male Mouse BAC"

ORIGIN
Query Match 6.2%; Score 56; DB 2; Length 190669;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 769 TCCTATCTGACGCGCTTCTGGGCGACTGAGTGGCGCCCTGTCAGAGCCCT 824
Db 177706 TCCATCTGACGCGCTTCTGGGCGACTGAGTGGCGCCCTGTCAGAGCCCT 177651

RESULT 28
AC079490 226909 bp DNA linear HTG 02-SEP-2000
LOCUS Mus musculus clone RP23-190A2, WORKING DRAFT SEQUENCE, 53 unordered
DEFINITION pieces.
AC079490
AC079490.1 GI:9964855
VERSION HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 226909)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Mouse
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 226909)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT -----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 1809834
Center clone name: RPCL-23_190A2

Summary Statistics
Consensus quality: 180161 bases at least Q40
Consensus quality: 198601 bases at least Q30
Consensus quality: 203733 bases at least Q20
Estimated insert size: 209000; agarose-fp estimation
Estimated insert size: 221709; sum-of-ctgigs estimation
Quality coverage: 6.2 in Q20 bases; agarose-fp estimation
Quality coverage: 5.84 in Q20 bases; sum-of-ctgigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1120: contig of 1120 bp in length

1121 1220: gap of unknown length
1221 2574: contig of 1354 bp in length
1251 2674: gap of unknown length
2675 4062: contig of 1388 bp in length
4063 4162: gap of unknown length
4163 5666: contig of 1504 bp in length
5667 5766: gap of unknown length
5767 6895: contig of 1129 bp in length
6896 6995: gap of unknown length
6996 8025: contig of 1030 bp in length
8026 8125: gap of unknown length
8126 9388: contig of 1263 bp in length
9389 9488: gap of unknown length
9489 10710: contig of 1222 bp in length
10711 10810: gap of unknown length
10811 11811: contig of 1001 bp in length
11812 11911: gap of unknown length
11912 12965: contig of 1054 bp in length
12966 13065: gap of unknown length
13066 14268: contig of 1203 bp in length
14269 14368: gap of unknown length
14369 15846: contig of 1478 bp in length
15847 15946: gap of unknown length
15947 17239: contig of 1273 bp in length
17240 17319: gap of unknown length
17320 18672: contig of 1353 bp in length
18673 18772: gap of unknown length
18773 19999: contig of 1227 bp in length
20000 20099: gap of unknown length
20100 21459: contig of 1360 bp in length
21460 21559: gap of unknown length
21560 22624: contig of 1063 bp in length
22625 22722: gap of unknown length
22723 23726: contig of 1004 bp in length
23727 23825: gap of unknown length
23826 25149: contig of 1323 bp in length
25150 25249: gap of unknown length
25250 26839: contig of 1590 bp in length
26840 26939: gap of unknown length
26940 28630: contig of 1691 bp in length
28631 28730: gap of unknown length
28731 30265: contig of 1535 bp in length
30266 30365: gap of unknown length
30366 31703: contig of 1338 bp in length
31704 31803: gap of unknown length
31804 32951: contig of 1148 bp in length
32952 33051: gap of unknown length
33052 35329: contig of 2278 bp in length
35330 35429: gap of unknown length
35430 38496: contig of 3067 bp in length
38497 38596: gap of unknown length
38597 40521: contig of 1925 bp in length
40522 40621: gap of unknown length
40622 42212: contig of 1591 bp in length
42213 42312: gap of unknown length
42314 44441: contig of 2129 bp in length
44442 44541: gap of unknown length
44542 46397: contig of 1856 bp in length
46398 46497: gap of unknown length
46499 48685: contig of 2188 bp in length
48686 48785: gap of unknown length
48786 50811: contig of 2026 bp in length
50812 50911: gap of unknown length
50912 52640: contig of 1729 bp in length
52641 52740: gap of unknown length
52741 53952: contig of 1212 bp in length
53953 54052: gap of unknown length
54053 58018: contig of 3966 bp in length
58019 58118: gap of unknown length
58119 61436: contig of 3318 bp in length
61437 61536: gap of unknown length
61537 64032: contig of 2496 bp in length
64033 64132: gap of unknown length

```

* 64133 66752: contig of 2620 bp in length
* 66753 66852: gap of unknown length
* 66853 70097: contig of 3245 bp in length
* 70098 70197: gap of unknown length
* 70198 73503: contig of 3306 bp in length
* 73504 73603: gap of unknown length
* 73604 80346: contig of 6743 bp in length
* 80347 80446: gap of unknown length
* 80447 86466: contig of 6020 bp in length
* 86467 86566: gap of unknown length
* 86567 92278: contig of 5712 bp in length
* 92279 92378: gap of unknown length
* 92379 98874: contig of 6496 bp in length
* 98875 98974: gap of unknown length
* 98975 106050: contig of 7076 bp in length
* 106051 106150: gap of unknown length
* 106151 113625: contig of 7475 bp in length
* 113626 113725: gap of unknown length
* 113726 129075: contig of 15350 bp in length
* 129076 129175: gap of unknown length
* 129176 141800: contig of 12625 bp in length
* 141801 141900: gap of unknown length
* 141901 155581: contig of 13681 bp in length
* 155582 155681: gap of unknown length
* 155682 170324: contig of 14643 bp in length
* 170325 170424: gap of unknown length
* 170425 184681: contig of 14257 bp in length
* 184682 184781: gap of unknown length
* 184782 206159: contig of 21378 bp in length
* 206160 206259: gap of unknown length
* 206260 226909: contig of 20650 bp in length.
location/Qualifiers
  1..226909
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /clone="RP23-190A2"
    /clone_1db="RPC1 mouse BAC library 23"

```

ORIGIN

```

Query Match      6.2%; Score 56; DB 2; Length 226909;
Best Local Similarity 100.0%; Pred. No. 1.5e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      769  TCCATCTGAGCGCCTCTGAGGAGGAGCTGAGTGAGCGCCCTGTCGACAGCCCT 824
Db      139528 TCCATCTGAGCGCCTCTGAGGAGGAGCTGAGTGAGCGCCCTGTCGACAGCCCT 139583

```

```

RESULT 29
BC075935      1821 bp      mRNA      linear      VRT 27-JUL-2004
LOCUS      Danio rerio zgc:92202, mRNA (CDNA clone MGC:92202 IMAGE:7051835),
DEFINITION      complete cds.
ACCESSION      BC075935
VERSION      BC075935.1 GI:49904383
KEYWORDS      MGC.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 1821)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Sherman, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapellato, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

```

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huik, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fallet, J., Helton, E., Kelleman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schermer, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1821)

Straussberg, R.

Direct Submission

Submitted (06-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: Len Zon, Harvard

cDNA Library Preparation: Open Biosystems

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES

source

```

1..1821
  /organism="Danio rerio"
  /mol_type="mRNA"
  /db_xref="taxon:7955"
  /clone="MGC:92202 IMAGE:7051835"
  /tissue_type="whole, adult zebrafish"
  /clone_1db="NIH ZGC_7"
  /lab_host="DH10B"
  /note="Vector: pExpress1"
  1..1821
    /gene="zgc:92202"
    /db_xref="locusID:436912"
    268..1482
      /gene="zgc:92202"
      /codon_start=1
      /product="zgc:92202"
      /protein_id="AA075935.1"
      /db_xref="GI:49904384"
      /db_xref="locusID:436912"
      /translation="MTSQHGNNAPALLSPONSSSGQRHISRPSIDYSRSLSSP
      RRGWVAASSGTHGNSLALSLRFLAPSSSSSTRPSPGVRPEPEAVDADYGL
      YSLRMPDIVGAQLTRHDVSLFLFDVIDEYRGIRSGRDFLLALERGRCDETN
      FRHYLQGLRIITRHDLLPYVTLRQGVCPDPVLYEISRVYSRGTVADACGTP
      HRRTGPOLICPEPSGPOVPRAPAPAPPRRRKRSHTTADCRQCTDIRLVA
      EYCHESALOGNVPSKQELEROFERFNQNTIKSRDAGSTICDIPSEITVLDAP
      WRDTISLLEALKGVFTTSLKQAVGHEAIKLVNDEEDYQGRKKLRNLVAGGA
      GAGTGSRESPLS"

```

gene

CDS

ORIGIN

```

Query Match      2.5%; Score 23; DB 5; Length 1821;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      745  TGTGACATCAAGTTCAGAGCT 767

```

Db 1228 TGTGACATCAAGTTCTCAGAGCT 1250

RESULT 30
LOCUS BX248385/c 125523 bp DNA linear VRT 19-NOV-2003
DEFINITION Zebrafish DNA sequence from clone CH211-229N2 in linkage group 8,
complete sequence.
ACCESSION BX248385
VERSION BX248385.6 GI:38453664
KEYWORDS HTG
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 125523)
Pelam, S.
Direct Submission
Submitted (19-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 19, 2003 this sequence version replaced gi:38304102.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived
Zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhifeng Bao and Sean Eddy, submitted), and those
beginning 'dir' were identified by Rick Waterman (Stephen Johnson
lab, WashU). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-229N2
is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.
Location/Qualifiers
1..125523
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-229N2"
/clone_id="CHORI-211"

ORIGIN
Query Match 2.5%; Score 23; DB 5; Length 125523;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 745 TGTGACATCAAGTTCTCAGAGCT 767

Db 63077 TGTGACATCAAGTTCTCAGAGCT 63055

RESULT 31
LOCUS CQ411697/c 838 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 18768 from Patent WO0170979.
ACCESSION CQ411697
VERSION CQ411697.1 GI:41319478
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Lee, J. and Lillie, J.
Genes, compositions, kits, and method for identification,
assessment, prevention, and therapy of ovarian cancer
Patent: WO 0170979-A 18768 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..838
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.4%; Score 22; DB 6; Length 838;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GGGAGAGAGATGAGTGCCTGGA 56

Db 381 GGGAGAGAGATGAGTGCCTGGA 360

RESULT 32
LOCUS AP001976/c 141272 bp DNA linear PRI 23-JAN-2001
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-179A16.
ACCESSION AP001976
VERSION AP001976.4 GI:12381917
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 141272)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 141,272 genomic DNA of 11q
2 (bases 1 to 141272)
Published Only in Database (2000)
1
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (02-MAY-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-cho, Tsukuba, Ibaraki, Japan
[E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170]
On Jan 22, 2001 this sequence version replaced gi:9927283.
STS Markers (ePCR):
22543 22694 D11S2068 11 Homo sapiens.
Location/Qualifiers
1..141272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"

FEATURES
SOURCE
Location/Qualifiers
1..141272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"


```

ORIGIN
      /clone="RP11-179A16"

Query Match      2.4%; Score 22; DB 9; Length 141272;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      558 CAGACCTTCTCTGAAGCAAA 579
      |||||
Db      90540 CAGACCTTCTCTGAAGCAAA 90519

RESULT 33
AP002382/c
LOCUS      154937 bp DNA linear HTG 31-MAY-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-50N19 map 11q, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
ACCESSION      AP002382.1 GI:8131646
VERSION      HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
      Fujiyama,A., Yada,T., Toroki,Y., Watanabe,H. and Sakaki,Y.
      Homo sapiens 154,937 genomic DNA of 11q
      Published only in Database (2000)
      2 (bases 1 to 154937)
      Direct Submission
      Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
      and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
      Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
      Japan (E-mail:hattori@gsc.riken.go.jp,
      URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
      Fax:81-42-778-9924)

COMMENT
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdract11
Center clone name: RP11-50N19
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 150397 bases at least Q40
Consensus quality: 152740 bases at least Q20
Consensus quality: 153635 bases at least Q30
Insert size: 154037; sum-of-contigs
Quality coverage: 6.50x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
10 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.

1
34030 contig of 34030 bp in length
34131 61053 contig of 26923 bp in length
61154 83303 contig of 22150 bp in length
83404 83303 contig of 18255 bp in length
101759 101758 contig of 15167 bp in length
117026 116925 contig of 14640 bp in length
141291 147116 contig of 9425 bp in length
147217 152406 contig of 5826 bp in length
152507 154937 contig of 2431 bp in length
152507 154937 contig of 2431 bp in length.
Location/Qualifiers
1. 154937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-50N19"
1. 34030
/note="assembly_fragment"
34131. 61053
/note="assembly_fragment"
61154. 83303
/note="assembly_fragment clone_end:SP6 vector_side:right"
83404. 101658
/note="assembly_fragment"
101759. 116925
/note="assembly_fragment"
117026. 131665
/note="assembly_fragment"
131766. 141190
/note="assembly_fragment clone_end:17 vector_side:left"
141291. 147116
/note="assembly_fragment"
147217. 152406
/note="assembly_fragment"
152507. 154937
/note="assembly_fragment"

```

```

FEATURES
source
1. 154937
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-50N19"
1. 34030
/note="assembly_fragment"
34131. 61053
/note="assembly_fragment"
61154. 83303
/note="assembly_fragment clone_end:SP6 vector_side:right"
83404. 101658
/note="assembly_fragment"
101759. 116925
/note="assembly_fragment"
117026. 131665
/note="assembly_fragment"
131766. 141190
/note="assembly_fragment clone_end:17 vector_side:left"
141291. 147116
/note="assembly_fragment"
147217. 152406
/note="assembly_fragment"
152507. 154937
/note="assembly_fragment"

ORIGIN
Query Match      2.4%; Score 22; DB 2; Length 154937;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      558 CAGACCTTCTCTGAAGCAAA 579
      |||||
Db      129827 CAGACCTTCTCTGAAGCAAA 129806

RESULT 34
LOCUS      CQ406469 323 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 13540 from Patent WO01/0979.
ACCESSION      CQ406469
VERSION      CQ406469.1 GI:41314250

```

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Lee, J. and Little, J.
TITLE Genes, compositions, kits, and method for identification, assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 13540 27-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..323
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 2.3%; Score 21; DB 6; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 582 GACCTGACATCCGCTCCG 602
Db 279 GACCTGACATCCGCTCCG 299

RESULT 35
AK060228
LOCUS AK060228
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:001-002-G12, full insert sequence.
ACCESSION AK060228
VERSION AK060228.1 GI:32970246
KEYWORDS FLI CDNA; oligo-capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE
AUTHORS The Rice Full-length CDNA Consortium, National Institute of Agricultural Sciences Rice Full-length CDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Mikiura, J., Ikeda, R., Iehibiki, J., Kawamata, M., Yoshimura, A., Mura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Inoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oato, N., Ota, Y., Saito, M., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE
AUTHORS 2 (bases 1 to 1553)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,

Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oato, N., Oca, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kanonndai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: ekikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT This clone is one of the 28k full-length cDNA clones from japonica rice.
URL: http://cdna01.dna.affrc.go.jp/cDNA/NIAS Rice Full-length CDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Iehibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J., Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oato, N., Oca, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.
FEATURES
source location/Qualifiers
1..1553
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-002-G12"

ORIGIN
Query Match 2.3%; Score 21; DB 8; Length 1553;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 667 CAGCGCTGGCGGCGACGTC 687
Db 1239 CAGCGCTGGCGGCGACGTC 1259

RESULT 36
AP006840 35/c
WPCOMMENT
Sequence split into 36 fragments LOCUS AP006840 Accession AP006840
Fragment Name Begin End
AP006840_00 1 11000
AP006840_01 100001 210000
AP006840_02 200001 310000
AP006840_03 300001 410000

AP006840_04 40001 51000
 AP006840_05 50001 61000
 AP006840_06 60001 71000
 AP006840_07 70001 81000
 AP006840_08 80001 91000
 AP006840_09 90001 101000
 AP006840_10 100001 111000
 AP006840_11 110001 121000
 AP006840_12 120001 131000
 AP006840_13 130001 141000
 AP006840_14 140001 151000
 AP006840_15 150001 161000
 AP006840_16 160001 171000
 AP006840_17 170001 181000
 AP006840_18 180001 191000
 AP006840_19 190001 201000
 AP006840_20 200001 211000
 AP006840_21 210001 221000
 AP006840_22 220001 231000
 AP006840_23 230001 241000
 AP006840_24 240001 251000
 AP006840_25 250001 261000
 AP006840_26 260001 271000
 AP006840_27 270001 281000
 AP006840_28 280001 291000
 AP006840_29 290001 301000
 AP006840_30 300001 311000
 AP006840_31 310001 321000
 AP006840_32 320001 331000
 AP006840_33 330001 341000
 AP006840_34 340001 351000
 AP006840_35 350001 3566135
 Continuation (36 of 36) of AP006840 from base 3500001 (AP006840 Symbiobacterium thermophilum IAM 14863)

Query Match 2.3%; Score 21; DB 1; Length 66135;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 210 GGAGCTGGAGCGCGGGGCA 230
 Db 40125 GGAGCTGGAGCGCGGGGCA 40105

RESULT 37 AP006840_00 WPCOMMIT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	11000
AP006840_01	10001	21000
AP006840_02	20001	31000
AP006840_03	30001	41000
AP006840_04	40001	51000
AP006840_05	50001	61000
AP006840_06	60001	71000
AP006840_07	70001	81000
AP006840_08	80001	91000
AP006840_09	90001	101000
AP006840_10	100001	111000
AP006840_11	110001	121000
AP006840_12	120001	131000
AP006840_13	130001	141000
AP006840_14	140001	151000
AP006840_15	150001	161000
AP006840_16	160001	171000
AP006840_17	170001	181000
AP006840_18	180001	191000
AP006840_19	190001	201000
AP006840_20	200001	211000
AP006840_21	210001	221000
AP006840_22	220001	231000
AP006840_23	230001	241000
AP006840_24	240001	251000

AP006840_25 250001 261000
 AP006840_26 260001 271000
 AP006840_27 270001 281000
 AP006840_28 280001 291000
 AP006840_29 290001 301000
 AP006840_30 300001 311000
 AP006840_31 310001 321000
 AP006840_32 320001 331000
 AP006840_33 330001 341000
 AP006840_34 340001 351000
 AP006840_35 350001 3566135
 LOCUS AP006840 3566135 bp DNA circular BCT 02-SEP-2004
 DEFINITION Symbiobacterium thermophilum IAM 14863
 ACCESSION AP006840
 VERSION AP006840.1 GI:51854827
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 gene
 CDS
 gene
 CDS

Ueda, K., Yamashita, A., Iehikawa, J., Shimada, M., Watsui, T.,
 Morimura, K., Ikeda, H., Hattori, M., and Bepu, T.,
 Complete genome sequence of an uncultured bacterium Symbiobacterium
 thermophilum
 2 (bases 1 to 3566135)
 Hattori, M., Yamashita, A., Toh, H., Oshima, K. and Shiba, T.
 Direct Submission
 Submitted (20-APR-2004) Masahira Hattori, Kitasato Institute for
 Life Sciences; 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail: hattori@genome.is.kitasato-u.ac.jp,
 URL: http://genome.is.kitasato-u.ac.jp, Tel: 81-42-778-8194)
 This work was done in collaboration with Kenji Ueda, Masafumi
 Shimada, Tomo-o Watsui, Kohji Morimura, Teruniko Bepu (Nihon
 University), Jun Ishikawa (National Institute of Infectious
 Diseases) and Haruo Ikeda (Kitasato University), and supported by
 the Research for the Future Program, 21st century COE program and
 the Grant-in-aid for the Scientific Research of MEXT, Japan.

location/Qualifiers
 1..3566135
 /organism="Symbiobacterium thermophilum IAM 14863"
 /mol_type="genomic DNA"
 /strain="IAM14863"
 /db_xref="taxon:292459"
 1..1377
 /gene="dna"
 /locus_tag="STH1"
 1..1377
 /gene="dna"
 /locus_tag="STH1"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="Chromosomal replication initiation protein"
 /protein_id="BAD38986.1"
 /db_xref="GI:51854828"
 /translation="MKTELSEVWOCIDHLERHTSKAIEVMVSTSPVAVVEDTIV
 GVQSEFAKRMESRYAPLRLQALBDLTRKMQLKFTVANNVIEDPAEPDAPVADL
 PAGTSAPAAQONARLLGYINPKYTFETVVGNSRFAPAAALAAVETPARTVPLFY
 GGVLGTHLMHAIGHVLOHNPTAKAVVSTETFTMAIOKSTTAFONRYRV
 DVLIIIDIQLAGKEATQEEFYHTFNAIKOIVISDRPKETIPLDRLRFE
 MGLICIDPDLERTRAILRKQASEIOVDVTVNATNIEINIREGLRRVVA
 YANKIKRLPYDLAVQALKDILPVRKQITIKVNAEHVNIIMODEVVRSSA
 VAYPRQYAMTLARELIDSLPKTIGEEGSDHTTVIACEXIKADIOSDPAFHTIO
 LIARIRAE"
 1600..2706
 /locus_tag="STH2"
 1600..2706
 /locus_tag="STH2"
 /codon_start=1
 /evidence=not experimental
 /transl_table=1
 /product="DNA polymerase III beta subunit"

/protein_id="BAD38987.1"
/db_xref="GI:51854829"
/translation="MKILISQESLNLTLAVAVSTKNTIPVLSGYVLSARGSELTIRADMEIAIAVTPCEVISEGTVLPARYFADLIRIRIPYGNIELSDVHNHNVATLWGRSVDYHGFADQPTLPDYDGASAFVKOSLIRLPTTPAAGHDESPKLTGYLFGKARDLLTGRTGIRIAYAAENAOGSPFVILIPSRSLNEVRLITDDESOVNAVAVANOVFPDLGTRVIRSLREGQYDPVWMLVQHPPTBAVLPGAEMELBEATLITKDAVAVKISFTQDKVTITSTNPEVGQYEBIAVTSWMDPLDIGNARYLLDLFLKLEEKQRFQCSGRNPARLMPADSQFLVYVPLITL"
2718..2945
/locus_tag="STH3"
2718..2945
/locus_tag="STH3"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAD38988.1"
/db_xref="GI:51854830"
/translation="MAEIVTITHEVYTLDAFLKWTGVADTGQAKALIAAGLVHNGEPELIRGRKLRADDRVRLGEGQMLIRREG"
2948..4075
/gene="recF"
/locus_tag="STH4"
2948..4075
/gene="recF"
/locus_tag="STH4"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="DNA repair and genetic recombination protein"
/protein_id="BAD38989.1"
/db_xref="GI:51854831"
/translation="MYLSTQLGAFRNYSLTITHPGLNVLVDGNAQKNTLLEAIFLATGRSHRTSRDPMWQEGREBELLAARVRGTLELRCGLQTRKQKINGIAEKTARIVGSLAVVLFSPDLOILKGPSPGRRLDLSOISOTYHLHLMANYRIVARNTLKQPVIDEGIMAYYDQVETGQVLYRRAREARLSPIASRHYRLADREBELAYVQSGVGDGADLETVRRRLERLELRSERKQVTLVGPHRDVGFWAGDARLYASQGQRTALALAKLAEFMSSEIGEPVLLDDVASELDPHRRHYLLSAVREGVQSFICTDLEDLMVEMPADHRLFRVAGTVVLDGRGS"
4075..4326
/locus_tag="STH5"
4075..4326
/locus_tag="STH5"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAD38990.1"
/db_xref="GI:51854832"
/translation="MYLHVGADVVALROYAIVNLPTVRKGGPTQELDLKRENRLLIOVEGGEARSILVLTADGAILSPISATTLKRSESAVLG"
4461..6377
/gene="gyrB"
/locus_tag="STH6"
4461..6377
/gene="gyrB"
/locus_tag="STH6"
/note="topoisomerase II"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="DNA gyrase B subunit"
/protein_id="BAD38991.1"
/db_xref="GI:51854833"
/translation="MNTHGTBOYTGEOIOVLGELAVRRBRGMYIGTSORGLLHLITWETIDNADBERLAGCDKIIVTTHSDGSVSDNRGIPVDIHPTGTISTEPTVYTLMAAGKFGGAYKVSGLHGVAAVNALSRKYLIOVKONGGVTEVEFANGGKTLPLKRVGTKEHGTYYRFPDPEIFESVDFDEAIVTRLDPAFLNQGLFIQIDEKAAEYRFVNEGLOTFVOYLNRSEPLHAKPIYVFKAIDTOEMELAIQYDAYSENILSFANNIRIVGGTHENGKLTALTRVINSYARKYLLKDNEDNLSGEDIREGMTAILISVKQESDPEQGTUKLGSSSEMTFVAAVSRGLSDPLEENPAIKRIYERKLSAARBEARKARBLVRKSALEITLALPGKLTDCSTRDPSSESLFVEGSGAGSAMGRDRRIQAVLPLR

GTIANEKARLDKILSNEIERLTLTAVGAGIDGDDPDISKORVHKVILIMTDADVDSHILRTLLTFPPRFRKPLIEACVLYIASPLVMIRKGVKQVYAYSQGEQDQYIARIGVTDKQGRVYGLDEBNMPEQLEMTTMDPSKILYKVTIMEDAAEATLFTLMGKEVPRREFIORVAHVQVMDLITG"
6411..8852
/gene="gyrA"
/locus_tag="STH7"
6411..8852
/gene="gyrA"
/locus_tag="STH7"
/note="topoisomerase II"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="DNA gyrase A subunit"
/protein_id="BAD38992.1"
/db_xref="GI:51854834"
/translation="MAGDFAORILITKEIKHEMETSYLTSMSVYVGRALPDVBDGLKPVORLILVAMYESGNTADKPPRSKAKTVGVLKSYHGHGDAAYDVMVMAQDFVLRPLVEGQNFSGSIDDPPPAARVTEBARLSRLAHEMLADIEKTVDFEDNYDNSEKEPTVLPSEFPNLVNGSTGIAVGMAVTNIPNNLGEVIDGILRLIDQPDISTEELNRIKGFDPFGALILIGHGIRRAYQVGKSYTLRKARQIETLQNGRRIIVTELPGVNAKGLIERIAELHRDKIDGITDILRDESDRGLIALELRDANPHYILNQLFVYALQCNFGIIMIALVNERPRLANLKMVHYLOHLEVIYARITFDLOKALDRAHILEGLLKALDIDIEVINTTISPTDDEARSRLMSNGFTEVQOHQITLDMQRLTLAREKLEVEYDELDRIAYTRAVLASERKQVYITIKQELGRIKKTGDPRTTITPDGIDIEDLIADBDVYV

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

cy 210 GGAGCTGGAGCGCGCGGCGCA 230
Db 32138 GGAGCTGGAGCGCGCGGCGCA 32158

RESULT 38
AP006840_09
WPCOMMENT
Sequence split into 36 fragments LOCUS AP006840 Accession AP006840
Fragment Name Begin End
AP006840_00 1 110000
AP006840_01 100001 210000
AP006840_02 200001 310000
AP006840_03 300001 410000
AP006840_04 400001 510000
AP006840_05 500001 610000
AP006840_06 600001 710000
AP006840_07 700001 810000
AP006840_08 800001 910000
AP006840_09 900001 1010000
AP006840_10 1000001 1110000
AP006840_11 1100001 1210000
AP006840_12 1200001 1310000
AP006840_13 1300001 1410000
AP006840_14 1400001 1510000
AP006840_15 1500001 1610000
AP006840_16 1600001 1710000
AP006840_17 1700001 1810000
AP006840_18 1800001 1910000
AP006840_19 1900001 2010000
AP006840_20 200001 2110000
AP006840_21 2100001 2210000
AP006840_22 2200001 2310000
AP006840_23 2300001 2410000
AP006840_24 2400001 2510000
AP006840_25 2500001 2610000
AP006840_26 2600001 2710000
AP006840_27 2700001 2810000
AP006840_28 2800001 2910000
AP006840_29 2900001 3010000
AP006840_30 3000001 3110000
AP006840_31 3100001 3210000

AP006840_32 320001 3310000
AP006840_33 330001 3410000
AP006840_34 340001 3510000
AP006840_35 350001 356135
Continuation (10 of 36) of AP006840 from base 900001 (AP006840 *Symbiobacterium thermophilum*)

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GGAGCTGAGCGCGCGGCA 230
DB 59411 GGAGCTGAGCGCGCGGCA 59431

RESULT 39
AP006840_12
WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000
AP006840_22	2200001	2310000
AP006840_23	2300001	2410000
AP006840_24	2400001	2510000
AP006840_25	2500001	2610000
AP006840_26	2600001	2710000
AP006840_27	2700001	2810000
AP006840_28	2800001	2910000
AP006840_29	2900001	3010000
AP006840_30	3000001	3110000
AP006840_31	3100001	3210000
AP006840_32	3200001	3310000
AP006840_33	3300001	3410000
AP006840_34	3400001	3510000
AP006840_35	3500001	356135

Continuation (13 of 36) of AP006840 from base 1200001 (AP006840 *Symbiobacterium thermophilum*)

Query Match 2.3%; Score 21; DB 1; Length 110000;

Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GGAGCTGAGCGCGCGGCA 230
DB 56924 GGAGCTGAGCGCGCGGCA 56944

RESULT 40
AP006840_14
WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840
Fragment Name Begin End
AP006840_00 1 110000

AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000
AP006840_22	2200001	2310000
AP006840_23	2300001	2410000
AP006840_24	2400001	2510000
AP006840_25	2500001	2610000
AP006840_26	2600001	2710000
AP006840_27	2700001	2810000
AP006840_28	2800001	2910000
AP006840_29	2900001	3010000
AP006840_30	3000001	3110000
AP006840_31	3100001	3210000
AP006840_32	3200001	3310000
AP006840_33	3300001	3410000
AP006840_34	3400001	3510000
AP006840_35	3500001	356135

Continuation (15 of 36) of AP006840 from base 1400001 (AP006840 *Symbiobacterium thermophilum*)

Query Match 2.3%; Score 21; DB 1; Length 110000;

Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GGAGCTGAGCGCGCGGCA 230
DB 92836 GGAGCTGAGCGCGCGGCA 92856

RESULT 41
AP006840_16
WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000

AP006840_22 220001 231000
AP006840_23 230001 241000
AP006840_24 240001 251000
AP006840_25 250001 261000
AP006840_26 260001 271000
AP006840_27 270001 281000
AP006840_28 280001 291000
AP006840_29 290001 301000
AP006840_30 300001 311000
AP006840_31 310001 321000
AP006840_32 320001 331000
AP006840_33 330001 341000
AP006840_34 340001 351000
AP006840_35 350001 356135
Continuation (17 of 36) of AP006840 from base 160001 (AP006840 *Symbiobacterium thermophilum*)

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 210 GGAGCTGGAGCGCGCGGCA 230
Db 86642 GGAGCTGGAGCGCGCGGCA 86662

RESULT 42
AP006840_18
WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000
AP006840_22	2200001	2310000
AP006840_23	2300001	2410000
AP006840_24	2400001	2510000
AP006840_25	2500001	2610000
AP006840_26	2600001	2710000
AP006840_27	2700001	2810000
AP006840_28	2800001	2910000
AP006840_29	2900001	3010000
AP006840_30	3000001	3110000
AP006840_31	3100001	3210000
AP006840_32	3200001	3310000
AP006840_33	3300001	3410000
AP006840_34	3400001	3510000
AP006840_35	3500001	356135

Continuation (19 of 36) of AP006840 from base 180001 (AP006840 *Symbiobacterium thermophilum*)

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 210 GGAGCTGGAGCGCGCGGCA 230

Db 13615 GGAGCTGGAGCGCGCGGCA 13635

RESULT 43
AP006840_21
WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000
AP006840_22	2200001	2310000
AP006840_23	2300001	2410000
AP006840_24	2400001	2510000
AP006840_25	2500001	2610000
AP006840_26	2600001	2710000
AP006840_27	2700001	2810000
AP006840_28	2800001	2910000
AP006840_29	2900001	3010000
AP006840_30	3000001	3110000
AP006840_31	3100001	3210000
AP006840_32	3200001	3310000
AP006840_33	3300001	3410000
AP006840_34	3400001	3510000
AP006840_35	3500001	356135

Continuation (22 of 36) of AP006840 from base 210001 (AP006840 *Symbiobacterium thermophilum*)

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 210 GGAGCTGGAGCGCGCGGCA 230
Db 56118 GGAGCTGGAGCGCGCGGCA 56138

RESULT 44
AP006840_21/c
WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000

AP006840_12 120001 1310000
AP006840_13 130001 1410000
AP006840_14 140001 1510000
AP006840_15 150001 1610000
AP006840_16 160001 1710000
AP006840_17 170001 1810000
AP006840_18 180001 1910000
AP006840_19 190001 2010000
AP006840_20 200001 2110000
AP006840_21 210001 2210000
AP006840_22 220001 2310000
AP006840_23 230001 2410000
AP006840_24 240001 2510000
AP006840_25 250001 2610000
AP006840_26 260001 2710000
AP006840_27 270001 2810000
AP006840_28 280001 2910000
AP006840_29 290001 3010000
AP006840_30 300001 3110000
AP006840_31 310001 3210000
AP006840_32 320001 3310000
AP006840_33 330001 3410000
AP006840_34 340001 3510000
AP006840_35 350001 3566135

Continuation (32 of 36) of AP006840 from base 2100001 (AP006840 Symbiodacterium thermoph

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GGAGCTGAGCGCCGCGGCA 230
DB 96361 GGAGCTGAGCGCCGCGGCA 96341

RESULT 45
AP006840_27/c

WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000
AP006840_22	2200001	2310000
AP006840_23	2300001	2410000
AP006840_24	2400001	2510000
AP006840_25	2500001	2610000
AP006840_26	2600001	2710000
AP006840_27	2700001	2810000
AP006840_28	2800001	2910000
AP006840_29	2900001	3010000
AP006840_30	3000001	3110000
AP006840_31	3100001	3210000
AP006840_32	3200001	3310000

AP006840_33 330001 3410000
AP006840_34 340001 3510000
AP006840_35 350001 3566135

Continuation (28 of 36) of AP006840 from base 2700001 (AP006840 Symbiodacterium thermoph

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GGAGCTGAGCGCCGCGGCA 230
DB 46788 GGAGCTGAGCGCCGCGGCA 46768

RESULT 46
AP006840_31/c

WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000
AP006840_02	200001	310000
AP006840_03	300001	410000
AP006840_04	400001	510000
AP006840_05	500001	610000
AP006840_06	600001	710000
AP006840_07	700001	810000
AP006840_08	800001	910000
AP006840_09	900001	1010000
AP006840_10	1000001	1110000
AP006840_11	1100001	1210000
AP006840_12	1200001	1310000
AP006840_13	1300001	1410000
AP006840_14	1400001	1510000
AP006840_15	1500001	1610000
AP006840_16	1600001	1710000
AP006840_17	1700001	1810000
AP006840_18	1800001	1910000
AP006840_19	1900001	2010000
AP006840_20	2000001	2110000
AP006840_21	2100001	2210000
AP006840_22	2200001	2310000
AP006840_23	2300001	2410000
AP006840_24	2400001	2510000
AP006840_25	2500001	2610000
AP006840_26	2600001	2710000
AP006840_27	2700001	2810000
AP006840_28	2800001	2910000
AP006840_29	2900001	3010000
AP006840_30	3000001	3110000
AP006840_31	3100001	3210000
AP006840_32	3200001	3310000
AP006840_33	3300001	3410000
AP006840_34	3400001	3510000
AP006840_35	3500001	3566135

Continuation (32 of 36) of AP006840 from base 3100001 (AP006840 Symbiodacterium thermoph

Query Match 2.3%; Score 21; DB 1; Length 110000;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 GGAGCTGAGCGCCGCGGCA 230
DB 52043 GGAGCTGAGCGCCGCGGCA 52023

RESULT 47
AP006840_33/c

WPCOMMENT

Sequence split into 36 fragments LOCUS AP006840 Accession AP006840

Fragment Name	Begin	End
AP006840_00	1	110000
AP006840_01	100001	210000

AP006840_02 200001 310000
 AP006840_03 300001 410000
 AP006840_04 400001 510000
 AP006840_05 500001 610000
 AP006840_06 600001 710000
 AP006840_07 700001 810000
 AP006840_08 800001 910000
 AP006840_09 900001 1010000
 AP006840_10 1000001 1110000
 AP006840_11 1100001 1210000
 AP006840_12 1200001 1310000
 AP006840_13 1300001 1410000
 AP006840_14 1400001 1510000
 AP006840_15 1500001 1610000
 AP006840_16 1600001 1710000
 AP006840_17 1700001 1810000
 AP006840_18 1800001 1910000
 AP006840_19 1900001 2010000
 AP006840_20 2000001 2110000
 AP006840_21 2100001 2210000
 AP006840_22 2200001 2310000
 AP006840_23 2300001 2410000
 AP006840_24 2400001 2510000
 AP006840_25 2500001 2610000
 AP006840_26 2600001 2710000
 AP006840_27 2700001 2810000
 AP006840_28 2800001 2910000
 AP006840_29 2900001 3010000
 AP006840_30 3000001 3110000
 AP006840_31 3100001 3210000
 AP006840_32 3200001 3310000
 AP006840_33 3300001 3410000
 AP006840_34 3400001 3510000
 AP006840_35 3500001 3566135

Continuation (34 of 36) of AP006840 from base 3300001 (AP006840 Symbiodacterium thermophilus)

Query Match 2.3%; Score 21; DB 1; Length 110000;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 210 GGAGCTGGAGCGCGCGCA 230

Db 21001 GGAGCTGGAGCGCGCGCA 20981

RESULT 48
 LOCUS AP005190/c 138141 bp DNA linear PLN 22-JUN-2004
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
 PAC clone: P0477A12.
 ACCESSION AP005190
 VERSION AP005190.3 GI:33146980
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Oryza sativa nuppondare (GA3) genomic DNA, chromosome 7, PAC
 clone: P0477A12
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 138141)
 AUTHORS Sasaki, T., Matsumoto, T. and Katayose, Y.
 TITLE Direct Submission
 JOURNAL Submitted (15-MAY-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 COMMENT On Jul 22, 2003 this sequence version replaced gi:24460086.
 Genes were predicted from the integrated results of the following:

FEATURES

source

GENSCAN (<http://CCR-081.mit.edu/GENSCAN.html>), FGENESH
 (<http://www.softberry.com/>), GeneMark.hmm
 (<http://opal.biology.gatech.edu/GeneMark/>), Glimmer
 (http://www.tigr.org/cdb/glimmer/glmr_form.html), RiceSHM
 (<http://rgp.dna.affrc.go.jp/RiceSHM/>), SplicePredictor
 (<http://bioinformatics.iastate.edu/cgi-bin/sp.cgi>), slm4
 (<http://globin.cse.psu.edu/html/docs/slm4.html>), gap2
 (<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI NonRedundant Protein
 database, nr (<http://ncbi.nlm.nih.gov/blast/db>) and the cDNA
 sequence database at RGP or DBP. Protein homologues of the coding
 regions were searched against NCBI NonRedundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DBP accession no. and RGP clone ID.
 BLASTX with the corresponding DBP accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from SP6 to T7 of the PAC clone.
 This sequence of P0477A12 clone has an overlap with OSJNB0064M11
 (DBP: AP005515) clone at 5' end and with OJ218.C12 (DBP:
 AP004051) clone at 3' end. Detailed information on overlap and
 assembly quality together with annotation of this entry is
 available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

Location/Qualifiers

1..138141

/organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nippondare"
 /db_xref="taxon:39947"
 /chromosome="7"
 /clone="P0477A12"
 complement(join(555..566,1727..1856,1980..2043,
 5143..5203))
 /gene="P0477A12.1"
 complement(join(555..566,1727..1856,1980..2043,
 5143..5203))
 /note="P0477A12.1"
 /note="hypothetical ORF
 predicted by GENSCAN
 this category is not included in IRGSP standard"
 complement(join(5913..6018,6085..6161))
 /gene="P0477A12.2"
 complement(join(5913..6018,6085..6161))
 /gene="P0477A12.2"
 /note="P0477A12.2"
 /note="hypothetical ORF
 predicted by Glimmer
 this category is not included in IRGSP standard"
 complement(6404..8115)
 /gene="P0477A12.3"
 complement(join(6404..7808,7968..8115))
 /gene="P0477A12.3"
 /note="supported by full-length cDNA(s): AK060228"
 complement(join(6554..7808,7968..8050))
 /gene="P0477A12.3"
 /note="contains full-length cDNA(s): AK060228"
 /codon_start=1
 /product="putative glucosyltransferase-12"
 /protein_id="BAC80053.1"
 /db_xref="GI:33146981"
 /translation="MAA...AAHVLFPFPAQGHNCMHFAGDLLEHESIRTKGSV
 AYRVLVASIVRAGDGSVGPVPTCVADIMPLAIDIAEIGEPALGPFVASC
 SFLAYISIRLREIGELPPRAGSDLDPEPRAGVGMTFLRRRLPBCRCGGGSDSN
 DPLQTVNEVTASRRARLVNTSASMSQFALAHIAFHRDVFALGFLHTMPAPAPA


```

gene
AGSLMRADGCMANLDGQPDNRVTVVSLGSEFTVISLQFTEFLHGLVAAGHAFIWIWR
PDWVGASQSAALREAAAAAGDSRARVWEMAPQLDLVLRHRAFCPLTHAGNMTLEAAG
EGVPTGVCMPFADQINSRFEVGYWRTGLDMKDVCDAAVVERWREAMESMETIASAQ
ALANOLRRDIADGSSAAEFQRLVGFITKELSGMSADSKSDLLTAGIRNALNN"
10498..11346
/gene="P0477A12.4"
/join(<10498..10638,10984..>11346)
/note="start and end point are not identified"
CDS
/join(10498..10638,10984..11346)
/gene="P0477A12.4"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC80054.1"
/db_xref="GI:33146982"
/translation="MGYVEDDKGYIEDDKGYVEESAHAOPAAVYANKKMYAKAOL
TRLFVLVGFQRCVTLCGLVSDGCVTAQFQSSRKRLRYITLTTSOPLTIKIV
IPNRGSAISQDGSSESRRLVLTLTWHSPTLASIKLVILLLAASVPLVFRTLL
SVNLPLP"
13200..13397
/gene="P0477A12.5"
13200..13397
/gene="P0477A12.5"
/note="hypothetical ORF
predicted by GlimmerX
this category is not included in IRGSP standard"
gene
/gene="P0477A12.6"
complement(join(<16265..16385,16509..16716,16914..>17016))
/note="start and end point are not identified"
CDS
complement(join(16265..16385,16509..16716,16914..17016))
/gene="P0477A12.6"
/note="contains EST(s) : AU163656(B4224)"
/codon_start=1
/product="putative ubiquitin / ribosomal protein CEP52"
/protein_id="BAC80055.1"
/db_xref="GI:33146983"
/translation="MGI FVKLTGKTTLTLEVSSDITIDNVAKIIOKEGIPPDQORLI
FAGQLEDGRTLAANYIOKESTLHLVLRRCGGPGPMIFPNLLALARSMEKMCIC
RKYARLPYKATNRKKKCGHTNEVILSILFEVHPDSGK"
gene
complement(join(18238..18435,22209..25030,27486..27685))
/gene="P0477A12.8"
complement(18238..18435)
/gene="P0477A12.8"
/note="3' terminal repeat"
/rpt_type="terminal"
18768..21827
/gene="P0477A12.7"
18768..21827
/gene="P0477A12.7"
/note="probably inactive due to including stop codon in
CDS
pseudogene, gag-pol precursor"
/misc_feature
complement(22209..25030)
/gene="P0477A12.8"
/note="probably inactive due to including frameshift in
CDS
pseudogene, Hypothetical protein with similarity to
putative retroelement"
/pseudo
/join(25422..25477,27168..27390)
/gene="P0477A12.9"
/join(<25422..25477,27168..>27390)
/note="start and end point are not identified"
CDS
/join(25422..25477,27168..27390)
/gene="P0477A12.9"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"

```

```

repeat_unit
/protein_id="BAC80056.1"
/db_xref="GI:33146984"
/translation="MKLPGRLEGRGNRRRERGAASIRCTAVGSLAAPKXKLVNGRV
SIEPRADGVGEBCDMFDPYVGFRCSENGRPDDPYTKMWRSLRCPH"
complement(27486..27685)
/gene="P0477A12.8"
/note="5' terminal repeat"
/rpt_type="terminal"
28892..29083
/gene="P0477A12.10"
<28892..>29083
/gene="P0477A12.10"
/note="start and end point are not identified"
CDS
28892..29083
/gene="P0477A12.10"
/note="predicted by GlimmerX etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAC80057.1"
/db_xref="GI:33146985"
/translation="MMARLTQIKRLASQGLGSFQNRARKEAWMLVHSSRAELSE
IESPSPRAHKPKLFFHPYT"
30837..32422
/gene="P0477A12.11"
/join(30837..31186,31300..31410,31513..31569,31655..31717,
31842..31904,32141..32422)
/gene="P0477A12.11"
/note="supported by full-length cDNA(s) : AK068191"

```

Query Match 2.3%; Score 21; DB 8; Length 138141;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Cy 667 CAGGCGCTGGCGGCGGACGCTG 687
Db 6718 CAGGCGCTGGCGGCGGACGCTG 6698

```

RESULT 49
LOCUS AP005515/c 164263 bp DNA linear PLN 22-JUL-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
ACCESSION AP005515
VERSION AP005515
KEYWORDS AP005515.3 GI:50508764
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
1 Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
clone:OSUNBA0064M1
Published Only in Database (2002)
2 (bases 1 to 164263)
JOURNAL Sasaki, T., Matsumoto, T. and Katayose, Y.
AUTHORS Direct Submission
TITLE Submitted (10-JUL-2002) Takuji Sasaki, National Institute of
JOURNAL Agrobiological Sciences, Rice Genome Research Program; Kannonnai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(8-mail:tsasa@nias.affrc.go.jp, URL:http://ryg.dna.affrc.go.jp/
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jul 22, 2004 this sequence version replaced gi:32490495.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), Glimmer
(http://www.tigr.org/cdb/glimmer/glmr_form.html), RiceNNM
(http://ryg.dna.affrc.go.jp/RiceNNM/), SplicePredictor
(http://bioinformatics.jaistate.edu/cgi-bin/ep.cgi), sim4
(http://glojin.cse.psu.edu/html/docs/sim4.html), gap2

(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologues of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRGSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from M13rev to -21M13 of the BAC clone. This sequence of OSJNBa0064M1.1 clone has an overlap with P0683C09 clone (DBJ: AP004348) at 5' end and an overlap with P0477A12 (DBJ: AP005190) at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomesSeq.html>.

FEATURES
source
1. 164263
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="7"
/clone="OSJNBa0064M1.1"
2337..4166
/gene="OSJNBa0064M1.1"
2337..4166
/gene="OSJNBa0064M1.1"
2337..4166
/note="probably inactive due to including stop codon in CDS"

gene
misc_feature
/gene="OSJNBa0064M1.1"
2337..4166
/note="probably inactive due to including stop codon in CDS"

gene
mRNA
CDS
/pseudogene, mdrb protein"
/complement (5307..5765)
/gene="OSJNBa0064M1.2"
/complement (<5307..5765)
/gene="OSJNBa0064M1.2"
/note="start and end point are not identified"
complement (5307..5765)
/gene="OSJNBa0064M1.2"
/note="predicted by GENESH etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD31524.1"
/db_xref="GI:50508765"
/translation="MAAMRRGSLATTNFEHPFPNPPLSFPPKQAGACLYMT
EDLRVEVEIVAYCHAGBFDPMETDELRWLDQGNRIELCSLLEKEVLOLY
RDELNOSKREVAEKHMESEKAMLQVVLVLMFICVLCVSVVGNKYLQ"
5794..5988
/gene="OSJNBa0064M1.3"
5794..5988
/gene="OSJNBa0064M1.3"
5794..5988
/note="start and end point are not identified"

gene
mRNA
CDS
/gene="OSJNBa0064M1.3"
5794..5988
/note="start and end point are not identified"
5794..5988
/gene="OSJNBa0064M1.3"
/note="predicted by Glimmer etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD31525.1"
/db_xref="GI:50508766"
/translation="MCLKLPATRKHTSARGCRTLPGTAPLYTCSTKSPGSIQPSGS
LPSRPARGLTVAATAAG"
join(6730..6883,7468..7529,7837..7929)
/gene="OSJNBa0064M1.4"

misc_feature

join(6730..6883,7468..7529,7837..7929)
/gene="OSJNBa0064M1.4"
/note="hypothetical ORF
predicted by RiceHMM
this category is not included in IRGSP standard"
10436..11579
/gene="OSJNBa0064M1.5"
join(<10436..10596,11573..>11579)
/gene="OSJNBa0064M1.5"
/note="start and end point are not identified"
join(10436..10596,11573..11579)
/gene="OSJNBa0064M1.5"
/note="predicted by GENSCAN etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD31526.1"
/db_xref="GI:50508767"
/translation="MEKTRERRELELVLAHKKERAEVQARKLAAQRAKADGS
EAKRKRTFPL"

gene
mRNA
CDS
complement (12516..16569)
/gene="OSJNBa0064M1.6"
complement (join(<12516..12815,13083..13233,13837..14074,
14171..14381,14478..14656,14773..14871,15684..>16569))
/gene="OSJNBa0064M1.6"
/note="start and end point are not identified"
complement (join(12516..12815,13083..13233,13837..14074,
14171..14381,14478..14656,14773..14871,15684..16569))
/gene="OSJNBa0064M1.6"
/codon_start=1
/product="receptor-like protein kinase-like protein"
/protein_id="BAD31527.1"
/db_xref="GI:50508768"
/translation="MALITVLPVLVLLPLLAIASGDPWQHCGSSGNTAGSKY
QANLQALAAITPSTASSSPALFAKDAAGGDAEPVAPLTICRDTASANASSSC
ADCAARFADQGVCPYSKVAAYYPCLLFSGDDPLSPAPADVRLVDVSTR
GGGADPFTLVRLALSYTQMAVPAVNSTGGAAYVWYTVRWYVTPPLFSIMCPT
DMSGDCRCQCLQDVNTTPNGSVSVRNIGARGCYRTPTKRYGSEPTLKGSLS
NSTAPSPPEPPVETRTSGRKVLTVALVPLALCPVATFCFPAITRIIRNKNMLR
KQDTMARREVLKMLREBSDESEFMPDFQIIBDATNFEEDKLGSGGFGSVYKQLP
NGLEVAIVRLAAHSQGLVEFKNEIOLIAKLTHTNLVNRGCCIOGSENNLIVYEPN
KSLDFPIPLKRAALLNWKTRLNIIIEGIIQGLLYLHGSHLCIIRHDLKASNTLLRD
MNPKISDFGLAKTIPDSNDVQNRKTRGVWAPYASGCFSLKSDVSPFGLVLE
ITSGKRNAGFHQYGDFFNLIGYAMQAKGSGWHELYDPSVSGCMEIKKCKVALL
CQENAVDPRTPMSAVVKKLSSEIKLIPKQPAFFNVRVHGGELNSTAPSSINDVIT
IVNGR"

gene
misc_feature
17641..18168
/gene="OSJNBa0064M1.7"
17641..18168
/gene="OSJNBa0064M1.7"
/note="hypothetical ORF
predicted by Glimmer"
this category is not included in IRGSP standard"
22180..22569
/gene="OSJNBa0064M1.8"
22180..22569
/gene="OSJNBa0064M1.8"
/note="start and end point are not identified"

gene
mRNA
CDS
22180..22569
/gene="OSJNBa0064M1.8"
/note="start and end point are not identified"

gene
mRNA
CDS
/gene="OSJNBa0064M1.8"
/note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD31528.1"
/db_xref="GI:50508769"
/translation="MTAGTRSGCGGDA PSPGSGRFGNATLGRAGSALGSSSVYGGG
FGARRQARVDPTTYCNQRAESEGGRHGGWGHGHPWTWIOQGVPTIRMRQRL
VPSRIWHRRRRRRTGTGDSFESAAMI"
complement (30486..30749)
/gene="OSJNBa0064M1.9"
complement (<30486..>30749)
/gene="OSJNBa0064M1.9"
/note="start and end point are not identified"
complement (30486..30749)

```

/gene="OSJNBa0064M11.9"
/note="predicted by FGENESH etc."
/codon_start=1
/product="hypothetical protein"
/db_xref="GI:50508770"
/translation="MSSLSGDKENVIPAKVEDISEQKAMEQKMOELKKAYINSFS
MTRQGYIOKXKCMPLKEGSGAGSGAQCEQEDNNKQMLWTL"
complement(35337..40496)
/gene="OSJNBa0064M11.10"
complement(35337..36095)
/gene="OSJNBa0064M11.10"
/note="3' terminal repeat"
/rpt_type="terminal"
complement(36116..40496)
/gene="OSJNBa0064M11.10"
/note="probably inactive due to including frameshift(s) in
CDS
probably inactive due to including stop codon(s) in CDS
pseudogene, polypeptide"
/pseudo

misc_feature

repeat_unit

gene

repeat_unit

misc_feature

Query Match 2.3%; Score 21; DB 8; Length 164263;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 667 CAGGCGCTGCGCGGCGAGCTG 687
Db 149571 CAGGCGCTGCGCGGCGAGCTG 149551

RESULT 50
AC117545/c 250084 bp DNA linear HTG 22-MAR-2004
DEFINITION Mus musculus chromosome 13 clone RP23-114J15 map 13, *** SEQUENCING
IN PROGRESS ***, 4 unordered pieces.
AC117545
AC117545.4 GI:45599005
HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 250084)
Birten,B., Nusbaum,C. and Lander,E.
Mus musculus chromosome 13, clone RP23-114J15
Unpublished
2 (bases 1 to 250084)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Lacroque,K., Lamazares,R.,
Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Melidiri,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riebeck,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Roselet,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE
Direct Submission
Submitted (10-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

```

REFERENCE
AUTHORS
3 (bases 1 to 250084)
Birten,B., Nusbaum,C., Lander,E., Aboueleil,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Lacroque,K., Lamazares,R.,
Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Melidiri,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riebeck,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Roselet,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE
Direct Submission
Submitted (22-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 22, 2004 this sequence version replaced gi:28913241.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT
JOURNAL
Submitted (22-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 22, 2004 this sequence version replaced gi:28913241.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

FEATURES
source
1. 250084
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chrname="13"
/map="13"
/clone="RP23-114J15"
/clone_1id="RP23-114J15"
Location/Qualifiers
1. 250084
65219 65318: contig of 65218 bp in length
65319 65318: gap of 100 bp
65319 154058: contig of 88740 bp in length
154059 154158: gap of 100 bp
154159 166023: contig of 11865 bp in length
166024 166123: gap of 100 bp
166124 250084: contig of 83961 bp in length.

ORIGIN
Query Match 2.3%; Score 21; DB 2; Length 250084;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 750 CATCAGTTCACAGCTCTC 770
Db 168962 CATCAGTTCACAGCTCTC 168942

Search completed: March 24, 2005, 00:19:34

```

Thu Mar 24 07:06:29 2005

us-10-030-271-1.oligo.rge

Page 34

Job time : 4106.48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 23, 2005, 16:37:44 ; Search time 3258.01 Seconds
(without alignments)
10620.110 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctaccgcggctgcac.....tcagctgcatgagctgcac 909

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

EST:
1: gb_esc1:.*
2: gb_esc2:.*
3: gb_hlc:.*
4: gb_esc3:.*
5: gb_esc4:.*
6: gb_esc5:.*
7: gb_esc6:.*
8: gb_gss1:.*
9: gb_gss2:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	858	94.4	1053	3	CR604307 full-leng
2	858	94.4	1860	3	CR593642 full-leng
3	858	94.4	1866	3	CR619301 full-leng
4	858	94.4	1894	3	CR625070 full-leng
5	607	66.8	1090	5	BM920838 AGENCOURT
6	586	64.5	800	4	BG685173 602637058
7	582	64.0	634	4	BI907024 603064980
8	582	64.0	905	4	BI910528 603068215
9	582	64.0	927	4	BG251151 602364854
10	582	64.0	947	4	BI910416 603068087
11	582	64.0	1062	5	BM908148 AGENCOURT
12	582	64.0	1131	2	BE907425 601500161
13	582	64.0	1510	3	CR599685 full-leng
14	582	64.0	1586	3	CR590827 full-leng
15	558	61.4	1089	5	BM925899 AGENCOURT
16	532	58.5	655	7	CK429257 o334b12.y
17	531	58.4	644	7	CN373290 170004241
18	531	58.4	993	4	BM559141 AGENCOURT
19	531	58.4	1051	5	BO073197 AGENCOURT
20	531	58.4	1161	5	BU902131 AGENCOURT
21	521	57.3	569	5	BP306442 BP306442
22	515	56.7	793	3	BE883568 601508040
23	508	55.9	654	2	BE391120 601286078
24	506	55.7	506	2	BF665646 602119860

25	502	55.2	678	7	CN373294
26	500	55.0	931	5	BQ923358
27	480	52.8	783	5	BU602002
28	476	52.4	718	4	BI117988
29	471	51.8	671	7	CN373287
30	458	50.4	577	7	CN373291
31	454	49.9	991	5	BQ956773
32	453	49.8	832	4	BI827768 603075430
33	450	49.5	1028	5	BM928759 AGENCOURT
34	450	49.5	1033	5	BM922715 AGENCOURT
35	446	49.1	586	5	BP343733
36	446	49.1	780	6	CB994488
37	442	48.6	831	6	CB992627
38	440	48.4	1042	5	BQ072022
39	435	47.9	931	5	BX381775
40	432	47.5	580	5	BP211545
41	429	47.2	583	5	BP305292
42	427	47.0	817	6	CB990003
43	426	46.9	891	5	BX458265
44	417	45.9	938	5	BU507654
45	415	45.7	1151	5	BQ947706
46	410	45.1	837	4	BI769587
47	407	44.8	1233	4	BM547754
48	403	44.3	582	5	BP253687
49	393	43.2	402	7	CN373296
50	388	42.7	475	2	AM245758
51	387	42.6	826	6	CB991766
52	385	42.4	582	5	BP360582
53	385	42.4	1428	5	BM906070
54	384	42.2	680	4	BG340752
55	383	42.1	576	4	BM126446
56	381	41.9	583	5	BP352691
57	380	41.8	495	7	CR746989
58	380	41.8	567	6	CB154807
59	378	41.6	527	4	BM837565
60	375	41.3	643	7	CN373297
61	372	40.9	586	5	BP254070
62	372	40.9	1313	5	BQ709748
63	368	40.5	657	7	CN373295
64	368	40.5	1043	5	BI180173
65	367	40.4	824	6	CB989607
66	362	39.8	583	5	BP304786
67	353	38.8	690	4	BG706385
68	349	38.4	574	7	CK819032
69	349	38.4	842	4	BI752583
70	344	37.8	581	5	BP333294
71	343	37.7	878	4	BI756130
72	340	37.4	474	7	CN373293
73	334	36.7	474	4	BM700298
74	325	35.8	415	7	CN373288
75	302	33.2	465	6	CB153442
76	300	33.0	799	7	BG681575
77	299	32.9	684	7	CN373289
78	297	32.7	926	4	BG757042
79	286	31.5	561	7	CN373292
80	285	31.4	546	4	BM849284
81	285	31.4	549	4	BM853352
82	285	31.4	623	4	BG819012
83	275	30.3	496	4	BG180215
84	268	29.5	522	4	BM147424
85	266	29.3	1046	5	BX342837
86	262	28.8	583	2	BE856624
87	262	28.8	587	5	BP306585
88	262	28.8	1481	5	BQ071492
89	254	27.9	421	6	CB153158
90	235	25.9	370	6	CB107743
91	235	25.9	872	4	BI753361
92	235	25.9	872	4	BI753389
93	230	25.3	823	2	BM837877
94	219	24.1	823	4	BE785785
95	216	23.8	855	2	BE797255
96	210	23.1	1021	5	BX384448
97	209	23.0	951	5	BX375652

CN373294	170005999
BQ923358	AGENCOURT
BU602002	AGENCOURT
BI117988	602867244
CN373287	170004251
CN373291	170004250
BQ956773	AGENCOURT
BI827768	603075430
BM928759	AGENCOURT
BM922715	AGENCOURT
BP343733	BP343733
CB994488	AGENCOURT
CB992627	AGENCOURT
BQ072022	AGENCOURT
BX381775	BX381775
BP211545	BP211545
BP305292	BP305292
CB990003	AGENCOURT
BX458265	BX458265
BU507654	AGENCOURT
BQ947706	AGENCOURT
BI769587	603054964
BM547754	AGENCOURT
BP253687	BP253687
CN373296	170005999
AM245758	2822785.5
CB991766	AGENCOURT
BP360582	BP360582
BM906070	AGENCOURT
BG340752	602462347
BM126446	11660609.y
BP352691	BP352691
CR746989	CR746989
CB154807	K-EST0212
BM837565	K-EST0113
CN373297	170006000
BP254070	BP254070
BQ709748	AGENCOURT
CN373295	170006000
BI180173	AGENCOURT
CB989607	AGENCOURT
BP304786	BP304786
BG706385	602669711
CK819032	11660609.y
BI752583	603028419
BP333294	BP333294
BI756130	603030029
CN373293	170006000
BM700298	UI-E-DW1
CN373288	170005326
CB153442	K-EST0210
BG681575	602628062
CN373289	170006000
BG757042	602710464
CN373292	170004705
BM849284	K-EST0129
BM853352	K-EST0134
BG819012	602781053
BG180215	602329620
BM147424	TCAPAP11
BX342837	BX342837
BE856624	765604.x
BP306585	BP306585
BQ071492	AGENCOURT
CB153158	K-EST0210
CB107743	K-EST0147
BI753361	603175851
BI753389	603029320
BM837877	K-EST0114
BE785785	601478567
BE797255	601587614
BX384448	BX384448
BX375652	BX375652

98	204	22.4	993	4	BG766140	BG766140	602738347
99	188	20.7	762	4	B1600659	B1600659	603247545
100	173	19.0	924	5	BX442929	BX442929	BX442929

ALIGNMENTS

RESULT 1	CR604307	1053 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	CR604307				
DEFINITION	full-length cDNA clone CSDDL006Y109 of B cells (Ramos cell line)				
ACCESSION	CR604307				
VERSION	CR604307.1	GI:50485114			
KEYWORDS	HTC; cNSLT cDNA.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1053)				
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
JOURNAL	Full-length cDNA libraries and normalization				
REMARK	Unpublished				
	Contact : Feng Liang Email : fliang@fietech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue				
REFERENCE	2 (bases 1 to 1053)				
AUTHORS	Genoscope.				
TITLE	Genoscope.				
JOURNAL	Direct Submission				
	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(AT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life technologies, a division of Invitrogen.				
FEATURES	Location/Qualifiers				
source	1..1053				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="CSDDL006Y109"				
	/tissue_type="B cells (Ramos cell line) Cot 25-normalized"				
	/plasmid="pCMVSPORT_6"				
ORIGIN					
	Query Match 94.4%, Score 858; DB 3; Length 1053;				
	Best Local Similarity 99.9%; Pred. Match: 0;				
	Matches 908; Conservative 1; Indels 1; Gaps 0;				
QY	1 ATGGCGCATCCGGGTGACCCCGGCGCGGTGCGGAGGAGATGAGTGCCTGAGTAC 60				
DB	99 ATGGGCTATCCGGTGCACCCCGGCGCGGTGCGGAGGAGATGAGTGCCTGAGTAC 158				
QY	61 TACGGAGTGTGTGCTTACCGTATGTTCAGGTGTGTGGCGGCACTGACCGAGTGC 120				
DB	159 TACGGAGTGTGTGCTTACCGTATGTTCAGGTGTGTGGCGGCACTGACCGAGTGC 218				
QY	121 GAGCTGAGCTTCGTGGCTTTTGTGTGATGAGGCTCTGTGGCGCGCGGAGGCTTACC 180				
DB	219 GAGCTGAGCTTCGTGGCTTTTGTGTGATGAGGCTCTGTGGCGCGCGGAGGCTTACC 278				
QY	181 CGGGCGCGCAGCGGCTGAGCTCTGTGTGAGCTGAGGCGCGCGGCACTGAGCGGAG 240				
DB	279 CGGGCGCGCAGCGGCTGAGCTCTGTGTGAGCTGAGGCGCGCGGCACTGAGCGAG 338				
QY	241 AGCAACTGCGGCTGTGGGCAACTCTGTGGCGGTGCTGGCGCGCAAGACTGTGTCGCG 300				
DB	339 AGCAACTGCGGCTGTGGGCAACTCTGTGGCGGTGCTGGCGCGCAAGACTGTGTCGCG 398				
QY	301 CACCTGCGCGCAAGCGCGCGCGCAAGTGTTCAGAACGCTATAGCTATGCACTTC 360				

Db	399	CACCTGGGGCCAAAGCGGCGCGCGACGTGCTCCAGAACGCTATAGCTATGGACCTCC	458
OY	361	AGCTTCTCAAGAGAGACAGAGGGTGAAGTCTGCGCGCTGGCGCTCAAGCAATTCTGCA	420
Db	459	AGCTTCTCAAGAGAGACAGAGGGTGAAGTCTGCGCGCTGGCGCTCAAGCAATTCTGCA	518
OY	421	AATTCTCAGCAGGGTCAAGTGGGAGACAGAGCTCCCCCACAACAAAGCGGACGGCGAGT	480
Db	519	AATTCTCAGCAGGGTCAAGTGGGAGACAGAGCTCCCCCACAACAAAGCGGACGGCGAGT	578
OY	481	CGGGGCGGCGCCCAAGTGGTGGTCCAGACAGCGCGGAGAGGGGCGCCACCGCACCCAG	540
Db	579	CGGGGCGGCGCCCAAGTGGTGGTCCAGACAGCGCGGAGAGGGGCGCCACCGCACCCAG	638
OY	541	CAGCAGTCAGAGCCCGCGACACTTCTCTGAAGGCAAGAGTACCTGTGACATCGGCTC	600
Db	639	CAGCAGTCAGAGCCCGCGACACTTCTCTGAAGGCAAGAGTACCTGTGACATCGGCTC	698
OY	601	CGGGTTCAGACAGATACCTGCGAGCATGGGCGAAGCTTTGAGACAGAGGCGTGGCATCCGG	660
Db	699	CGGGTTCAGACAGATACCTGCGAGCATGGGCGAGCTTTGAGACAGAGGCGTGGCATCCGG	758
OY	661	CGGCCCCAGGCGCTGGCGCGGACAGCTGGACGCTTTGGGACAGGCAACCGACAGTGTGGC	720
Db	759	CGGCCCCAGGCGCTGGCGCGGACAGCTGGACGCTTTGGGACAGGCAACCGACAGTGTGGC	818
OY	721	TCAAGGACCTGGGCTGTGTGTGTGTGATCAATCAAGTTCTAGAGCTCTCTATCTGAC	780
Db	819	TCAAGGACCTGGGCTGTGTGTGTGTGATCAATCAAGTTCTAGAGCTCTCTATCTGAC	878
OY	781	GCGTTCTGGGGGCACTATCTGATGGGCGCCTGCGTGCAGAGCGCTGGCGGGCGTGTCCG	840
Db	879	GCGTTCTGGGGGCACTATCTGATGGGCGCCTGCGTGCAGAGCGCTGGCGGGCGTGTCCG	938
OY	841	ACTAGGCGCTTCGAGAGAGCTGTGGGCGGAGAGCTGTTCGCTGCTGTGATGTGAT	900
Db	939	ACTAGGCGCTTCGAGAGAGCTGTGGGCGGAGAGCTGTTCGCTGCTGTGATGTGAT	998
OY	901	GAGGCTGAC 909	
Db	999	GAGGCTGAC 1007	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	JOURNAL	COMMENT
CR593642	full-length cDNA clone	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Unpublished			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/invitrogen Corporation 1600 Faraday Avenue			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	2 (bases 1 to 1860)			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Genoscope.			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Direct Submission			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Web : www.genoscope.cns.fr			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	end enriched, double-strand cDNA was digested with Not I and cloned			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library			
CR593642	1860 bp mRNA linear	CR593642	CR593642.1	HTC; CNSLT cDNA.	Homo sapiens (human)	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	was normalized. Library was constructed by Life Technologies, a			

```

FEATURES
  source
    division of Invitrogen.
    Location/Qualifiers
      1..1860
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DK010YK16"
        /tissue_type="Hela cells Cot 25-normalized"
        /plasmid="PCMVSPORT_6"
ORIGIN
Query Match      94.4%; Score 858; DB 3; Length 1860;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ATGGGCGCTATCCGGGTGACCCCGGCGCTGGGAGAGAGATGCTGGACTAC 60
Db 71 ATGGCGCTATCCGGGTGACCCCGGCGCTGGGAGAGAGATGCTGGACTAC 130
Oy 61 TACGGGATGCTGCTGCTTACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 120
Db 131 TACGGGATGCTGCTGCTTACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 190
Oy 121 GAGCTGAGCTTCCTGGCTTTTCTGCTGATGAGCTCTTGGGCGCGGAGGCTTAGCC 180
Db 191 GAGCTGAGCTTCCTGGCTTTTCTGCTGATGAGCTCTTGGGCGCGGAGGCTTAGCC 250
Oy 181 CGGGCGCGGAGCGGCTAGAGTCTCTGCTGAGCTGAGCGCGCGGCGGAGTGGCGGAG 240
Db 251 CGGGCGCGGAGCGGCTAGAGTCTCTGCTGAGCTGAGCGCGCGGCGGAGTGGCGGAG 310
Oy 241 AGCAACTGCGGCTGCTGCTGGGCAACTCCTGCGCGTGTGCGCGCGCACGACTGTGCGG 300
Db 311 AGCAACTGCGGCTGCTGCTGGGCAACTCCTGCGCGTGTGCGCGCGCACGACTGTGCGG 370
Oy 301 CACCTGCGCGCAAGCGGCGCGGCGAGTGTCTCAGAACGCTATAGCTTAGGCACTTCC 360
Db 371 CACCTGCGCGCAAGCGGCGCGGCGAGTGTCTCAGAACGCTATAGCTTAGGCACTTCC 430
Oy 361 AGCTCTCAAGAGGAGCAGAGGATGAGTCCGCTGCGCGCTCAAGAGCTTCTGCA 420
Db 431 AGCTCTTCAAGAGGAGCAGAGGATGAGTCCGCTGCGCGCTCAAGAGCTTCTGCA 490
Oy 421 AATTCTCAGCAGAGCTCAGTGGGAGACAGAGCTCCCCCAACCAAGCGGCGGAGAGT 480
Db 491 AATTCTCAGCAGAGCTCAGTGGGAGACAGAGCTCCCCCAACCAAGCGGCGGAGAGT 550
Oy 481 CGGGGCGCGGCGAGTGTGTGCTCCAGACGCGCGGAGAGGCGGCCCAAGCGGCAAG 540
Db 551 CGGGGCGCGGCGAGTGTGTGCTCCAGACGCGCGGAGAGGCGGCCCAAGCGGCAAG 610
Oy 541 CAGCAGTGAAGAGCCCGGCGAGACCTTCTCTGAGAGCAAGTGAACCTGTGACATCCGGCTC 600
Db 611 CAGCAGTGAAGAGCCCGGCGAGACCTTCTCTGAGAGCAAGTGAACCTGTGACATCCGGCTC 670
Oy 601 CGGGTTCAGAGAGAGTACTGCGAGATGAGGCGCTTGGAGCAGAGGAGGCTGAGTCCCGG 660
Db 671 CGGGTTCAGAGAGAGTACTGCGAGATGAGGCGCTTGGAGCAGAGGAGGCTGAGTCCCGG 730
Oy 661 CGGGCGCGGAGCGCTGCGCGGAGCTGAGCGTGTGGGCAAGGCCACGCAAGTGTGCGC 720
Db 731 CGGGCGCGGAGCGCTGCGCGGAGCTGAGCGTGTGGGCAAGGCCACGCAAGTGTGCGC 790
Oy 721 TCAAGGAGACTGGGCTGTGTGTTGATCAATCAAGTCTCAGAGCTTCTCATTTGAGAC 780
Db 791 TCAAGGAGACTGGGCTGTGTGTTGATCAATCAAGTCTTCTCATTTGAGAC 850
Oy 781 GCTTCTGGGCGAGTACTGAGTGGCGCCCTGCTGACAGGCGCTGCGGGGCGTGTTCCTG 840
Db 851 GCTTCTGGGCGAGTACTGAGTGGCGCCCTGCTGACAGGCGCTGCGGGGCGTGTTCCTG 910
Oy 841 ACTGAGGCTCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGCTGCTGAGTGTGAT 900

```

```

Db 911 ACTAGGCGCTTCGAGAGAGGCTGTGGGCGGAGGCTGTTGCTGCTGCTGATGATGAT 970
Oy 901 GAGGCTGAC 909
Db 971 GAGGCTGAC 979

RESULT 3
CR619301
LOCUS
DEFINITION
  full-length cDNA clone CS0DC015YJ24 of Neuroblastoma Cot
  25-normalized of Homo sapiens (human).
ACCESSION
  CR619301
VERSION
  CR619301.1 GI:50500108
KEYWORDS
  HTC; CNSLT cDNA.
SOURCE
  ORGANISM
    Homo sapiens (human)
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 1866)
  Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
  Full-length cDNA libraries and normalization
  Unpublished
  Contact : Feng Liang Email : fliang@lifetech.com URL :
  http://fulllength.invitrogen.com/Invitrogen Corporation 1600
  Faraday Avenue
  2 (bases 1 to 1866)
REFERENCE
  Genoscope.
  Direct Submission
  Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
  BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
  - Web : www.genoscope.cns.fr)
  1st strand cDNA was primed with a NotI-oligo(dT) primer. Five clones
  into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
  was normalized. Library was constructed by Life Technologies, a
  division of Invitrogen.
FEATURES
  source
    Location/Qualifiers
      1..1866
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="CS0DC015YJ24"
        /tissue_type="Neuroblastoma Cot 25-normalized"
        /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      94.4%; Score 858; DB 3; Length 1866;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ATGGCGCTATCCGGGTGACCCCGGCGCTGGGAGAGAGATGCTGGACTAC 60
Db 76 ATGGCGCTATCCGGGTGACCCCGGCGCTGGGAGAGAGATGCTGGACTAC 135
Oy 61 TACGGGATGCTGCTGCTTACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 120
Db 136 TACGGGATGCTGCTGCTTACCGTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 195
Oy 121 GAGCTGAGCTTCCTGGCTTTTCTGCTGATGAGCTCTTGGGCGCGGAGGCTTAGCC 180
Db 196 GAGCTGAGCTTCCTGGCTTTTCTGCTGATGAGCTCTTGGGCGCGGAGGCTTAGCC 255
Oy 181 CGGGCGCGGAGCGGCTTAGAGCTCTCTGCTGAGAGCTGAGCGCGCGGCACTGCGGAG 240
Db 196 GAGCTGAGAGCTCTGAGCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 255
Oy 181 CGGGCGCGGAGCGGCTTAGAGCTCTCTGCTGAGAGCTGAGCGCGCGGCACTGCGGAG 240
Db 256 CGGGCGCGGAGCGGCTTAGAGCTCTCTGCTGAGAGCTGAGCGCGCGGCACTGCGGAG 315
Oy 241 AGCAACTGCGGCTGCTGCTGGGCAACTCCTGCGCGTGTGCGCGCGCACGACTGTGCGG 300
Db 316 AGCAACTGCGGCTGCTGCTGGGCAACTCCTGCGCGTGTGCGCGCGCACGACTGTGCGG 375
Oy 301 CACCTGCGCGCAAGCGGCGCGGCGAGTGTCTCAGAACGCTATAGCTTAGGCACTTCC 360

```

Db 376 CACCTGGCGGCAAGCGGCGCCGCACTGCTCCAGACGCTATAGCTATGCACTCC 435
Qy 361 AGCTCTTCAAGAAGAGAGAGAGTGTAGTCCGCTGCGCCGTGCGAGTCAAGCACTTGTGA 420
Db 436 AGCTCTTCAAGAAGAGAGAGAGTGTAGTCCGCTGCGCCGTGCGAGTCAAGCACTTGTGA 495
Qy 421 AATTCTCAGACAGGTCTAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGAGCGGAGT 480
Db 496 AATTCTCAGACAGGTCTAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGAGCGGAGT 555
Qy 481 CCGGGCGCGCCAGTGTGTGTGTCAGACGCGCGGAGAGAGGGGCGCCAGCGCCAG 540
Db 556 CCGGGCGCGCCAGTGTGTGTGTCAGACGCGCGGAGAGAGGGGCGCCAGCGCCAG 615
Qy 541 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAAGCAAAAGTGAACCTGTGAATCCGCTC 600
Db 616 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAAGCAAAAGTGAACCTGTGAATCCGCTC 675
Qy 601 CCGGTTTGAGAGAGTACTGTGAGACATGGGCAACCTTTGAGAGCGGCGTGGCATCCCG 660
Db 676 CCGGTTTGAGAGAGTACTGTGAGACATGGGCAACCTTTGAGAGCGGCGTGGCATCCCG 735
Qy 661 CCGGCGCCAGGCGCTGCGCGCGGCACTGGAACGTGTTTGGGCGAGCCAGCTGCTGCG 720
Db 736 CCGGCGCCAGGCGCTGCGCGCGGCACTGGAACGTGTTTGGGCGAGCCAGCTGCTGCG 795
Qy 721 TCAAGGAGACCTGGGCTGTGTGTTGTGACATCAAGTTCTTGAGACTTCTTATCTGAG 780
Db 796 TCAAGGAGACCTGGGCTGTGTGTTGTGACATCAAGTTCTTGAGACTTCTTATCTGAG 855
Qy 781 GCTTTCTGGGGCGACTACCTGAGATGGGCGCTGCTGAGAGCGCTGGCGGGGCTTCTTG 840
Db 856 GCTTTCTGGGGCGACTACCTGAGATGGGCGCTGCTGAGAGCGCTGGCGGGGCTTCTTG 915
Qy 841 ACTGAGGCGCTGCGAGAGAGTGTGAGCGCGGAGGCGTGTGCTGCTGCTGCTGAGTGG 900
Db 916 ACTGAGGCGCTGCGAGAGAGTGTGAGCGCGGAGGCGTGTGCTGCTGCTGCTGAGTGG 975
Qy 901 GAGGCTGAC 909
Db 976 GAGGCTGAC 984

RESULT 4
CR625070 1894 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA CSODC001YC18 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
ACCESSION CR625070
VERSION CR625070.1 GI:50505877
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1894)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@life.technet.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue
2 (bases 1 to 1894)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: seqdef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(47) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
Source Location/Qualifiers
1..1894
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC001YC18"
/isolate="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 94.4%; Score 858; DB 3; Length 1894;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGAGATGCTCTGAGTAC 60
Db ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGAGATGCTCTGAGTAC 148
Qy 89
Db 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGATGTTGGCGGCACTGACCGAGTGC 120
Qy 149 TACGGATGCTGTGCTTCAACCGTATGTTGAGATGTTGGCGGCACTGACCGAGTGC 208
Db 121 GAGCTGAGCTCTGAGCTTCTGCTGATGATGAGCTCTGCGCCGCGAGGCTTAGCC 180
Qy 209 GAGCTGAGCTCTGAGCTTCTGCTGATGATGAGCTCTGCGCCGCGAGGCTTAGCC 268
Db 181 CCGGCGCGAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGCGGCACTGCGCGAG 240
Qy 269 CCGGCGCGAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGCGGCACTGCGCGAG 328
Db 241 AGCAACTGTGGGCTGTGCTGGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Qy 329 AGCAACTGTGGGCTGTGCTGGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 388
Db 301 CACTGCGCGCAAGCGGCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 389 CACTGCGCGCAAGCGGCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 448
Db 449 AGCTCTTCAAGAAGAGAGAGAGTGTAGTCCGCTGCGCGAGTCAACAGCTTGTGCA 508
Qy 421 AATTCTCAGACAGGTCTAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGAGCGGAGT 480
Db 509 AATTCTCAGACAGGTCTAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGAGCGGAGT 568
Qy 481 CCGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGAGGGGCGCCAGCGCACCCAG 540
Db 569 CCGGGCGCGCCAGTGTGTGTCAGACGCGCGGAGAGAGGGGCGCCAGCGCACCCAG 628
Qy 541 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAAGCAAAAGTGAACCTGTGAATCCGCTC 600
Db 629 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAAGCAAAAGTGAACCTGTGAATCCGCTC 688
Qy 601 CCGGTTTGAGAGAGTACTGTGAGACATGGGCAACCTTTGAGAGCGGCGTGGCATCCCG 660
Db 689 CCGGTTTGAGAGAGTACTGTGAGACATGGGCAACCTTTGAGAGCGGCGTGGCATCCCG 748
Qy 661 CCGGCGCCAGGCGCTGCGCGCGGCACTGGAACGTGTTTGGGCGAGCCAGCTGCTGCG 720
Db 749 CCGGCGCCAGGCGCTGCGCGCGGCACTGGAACGTGTTTGGGCGAGCCAGCTGCTGCG 808
Qy 721 TCAAGGAGACCTGGGCTGTGTGTTGTGACATCAAGTTCTTGAGACTTCTTATCTGAG 780
Db 809 TCAAGGAGACCTGGGCTGTGTGTTGTGACATCAAGTTCTTGAGACTTCTTATCTGAG 868
Qy 781 GCTTTCTGGGGCGACTACCTGAGATGGGCGCTGCTGAGAGCGCTGGCGGGGCTTCTTG 840
Db 869 GCTTTCTGGGGCGACTACCTGAGATGGGCGCTGCTGAGAGCGCTGGCGGGGCTTCTTG 928
Qy 841 ACTGAGGCGCTGCGAGAGAGTGTGAGCGCGGAGGCTGCTGCTGCTGCTGCTGAGTGG 900
Db 929 ACTGAGGCGCTGCGAGAGAGTGTGAGCGCGGAGGCTGCTGCTGCTGCTGAGTGG 988


```

Qy 901 GAGGCTGAC 909
Db 989 GAGGCTGAC 997

RESULT 5
BM920838
LOCUS
DEFINITION
  BM920838 1090 bp mRNA linear EST 12-MAR-2002
  AGENCOURT 6706034 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752092
  5', mRNA sequence.
ACCESSION
  BM920838
VERSION
  BM920838.1 GI:19371217
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Organism
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1 (bases 1 to 1090)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue Procurement: Life Technologies, Inc.
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
  DNA Sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LNLN at:
  http://image.llnl.gov
  Plate: L16M12785 row: 1 column: 13
  High quality sequence stop: 625.
FEATURES
  source
  1..1090
  location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:5752092"
  /lab_host="DH10B"
  /clone_id="NIH_MGC_115"
  /note="Organ: pooled brain, lung, testis; Vector:
  pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
  source anonymous pool of 6 male brains, age range 23-27; 1
  male lung, age 27; and 1 male testis, age 69. Library is
  oligo-dT primed and directionally cloned. (EcoRV site is
  destroyed upon cloning). Average insert size 1.8 Kb,
  insert size range 1-3 Kb. Library is normalized and
  enriched for full-length clones and was constructed by C.
  Gruber (Invitrogen). Research Genetics tracking code
  021. Note: this is a NIH_MGC library."

ORIGIN
Query Match 66.8%; Score 607; DB 5; Length 1090;
Best Local Similarity 99.8%; Pred. No. 2.2e-297;
Matches 657; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCATCCGGGTGACCCCGGCGCTGGGAGGATGATGCTGGACTAC 60
Db 102 ATGGCGCATCCGGGTGACCCCGGCGCTGGGAGGAGATGCTGGACTAC 161
Qy 61 TACGGGATGCTGCTGCTTCCATGATGTTGAGGTGAGGCGGCACTGACGAGTGC 120
Db 162 TACGGGATGCTGCTTCCATGATGTTGAGGTGAGGCGGCACTGACGAGTGC 221
Qy 121 GAGCTGAGCTCTGCTGCTTCTGCTGATAGAGCTCTGCGCGCGGAGGCTTACC 180
Db 222 GAGCTGAGCTCTGCTTCTGCTGATAGAGCTCTGCGCGCGGAGGCTTACC 281
Qy 181 CGGGCCGACGCGGCTGAGACTCTGCTGAGAGCTGAGCGCGCGGCACTGAGCGCAG 240
Db 282 CGGGCCGACGCGGCTGAGACTCTGCTGAGAGCTGAGCGCGCGGCACTGAGCGCAG 341

```

```

Qy 241 AGCAACTGCGGCTGCTGGGCAACTCTCTGCGCTGCTGCGCCGACGACTCTGCGG 300
Db 342 AGCAACTGCGGCTGCTGGGCAACTCTCTGCGCTGCTGCGCCGACGACTCTGCGG 401
Qy 301 CACCTGGCGCCAGCGCGCGCGCGCTGCTCTCCAGAACCTTATGATGCACTCC 360
Db 402 CACCTGGCGCCAGCGCGCGCGCGCTGCTCTCCAGAACCTTATGATGCACTCC 461
Qy 361 AGCTCTTCAAGAGAGACAGAGGATGAGCTGCGCGCGCGCGAGTCAAGGATTTGCA 420
Db 462 AGCTCTTCAAGAGAGACAGAGGATGAGCTGCGCGCGCGCGAGTCAAGGATTTGCA 521
Qy 421 AATTCTGACAGGCTGCTGAGAGACAGGCTCTCCCGCAACAGCGGACGCGGAGT 480
Db 522 AATTCTGACAGGCTGCTGAGAGACAGGCTCTCCCGCAACAGCGGACGCGGAGT 581
Qy 481 CGGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 582 CGGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 641
Qy 541 CAGAGTCAGAGCGCGCGCGCGCGCGCTTCTGTAAGGCAAGTGAATGATCGGCTC 600
Db 642 CAGAGTCAGAGCGCGCGCGCGCGCGCTTCTGTAAGGCAAGTGAATGATCGGCTC 701
Qy 601 CGGGTTGAGAGAGTACTGCTGAGCATGAGCGGCGGCTTGGAGAGCGGCTGCAATCC 658
Db 702 CGGGTTGAGAGAGTACTGCTGAGCATGAGCGGCGGCTTGGAGAGCGGCTGCAATCC 759

RESULT 6
BM685173
LOCUS
DEFINITION
  BM685173 800 bp mRNA linear EST 01-MAY-2001
  602637058P1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:474639 5',
  mRNA sequence.
ACCESSION
  BM685173
VERSION
  BM685173.1 GI:13916570
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Homo sapiens
  Organism
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1 (bases 1 to 800)
  NIH-MGC http://mgi.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgabbs-remail.nih.gov
  Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
  CDNA Library Preparation: Ling Hong/Rubin Laboratory
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LNLN at:
  http://image.llnl.gov
  Plate: L16M1622 row: d column: 16
  High quality sequence stop: 794.
FEATURES
  source
  1..800
  location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:474639"
  /tissue_type="Primary B-cells from tonsils (cell line)"
  /lab_host="DH10B (phage-resistant)"
  /clone_id="NIH_MGC_48"
  /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
  Site_2: EcoRI; cDNA made by oligo-dT priming.
  Directionally cloned into EcoRI/XhoI sites using the
  following 5' adaptor: GGCAGAG(G). Size-selected >500bp
  for average insert size 1.8kb library constructed by Ling
  Hong in the laboratory of Gerald M. Rubin (University of
  California, Berkeley) using ZAP-cDNA synthesis kit
  (Stratagene) and Superscript II RT (Life Technologies)."

```

ORIGIN Note: this is a NIH_MGC library."

Query Match 64.5%; Score 586; DB 4; Length 800;
Best Local Similarity 100.0%; Pred. No. 1.1e-286;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 CGAGAGCAACCTGCGCTGCTGGGGCAACTCTGCGGTGTCGCCGCCCAAGCACTGCT 296
DB 119 CGAGAGCAACCTGCGCTGCTGGGGCAACTCTGCGGTGTCGCCGCCCAAGCACTGCT 178
QY 297 GCGGCACTGCGCGGCAAGGCGCGCGGCAAGTCTCCAGAGCTATAGCTATGACAC 356
DB 179 GCGGCACTGCGCGGCAAGGCGCGCGGCAAGTCTCCAGAGCTATAGCTATGACAC 238
QY 357 CTCAGCTCTTCAAGAGGAGCAGAGGAGTCTGCGCTGCGCTGCGAGCTCAAGCACTTC 416
DB 239 CTCAGCTCTTCAAGAGGAGCAGAGGAGTCTGCGCTGCGCTGCGAGCTCAAGCACTTC 298
QY 417 TGCAGATTCAGCAGGCTCAGTGGAGACAGGCTCCCGCCCAAGCGGCGGCGG 476
DB 299 TGCAGATTCAGCAGGCTCAGTGGAGACAGGCTCCCGCCCAAGCGGCGGCGGCGG 358
QY 477 GAGTCGGGGCGCGCCAGTGGTGGTGCAGACGCGGCGGAGAGGCGGCCAGCGGACG 536
DB 359 GAGTCGGGGCGCGCCAGTGGTGGTGCAGACGCGGCGGAGAGGCGGCCAGCGGACG 418
QY 537 CCAAGCAGTCAAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCG 596
DB 419 CCAAGCAGTCAAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGAATCTGTGACATCCG 478
QY 597 GCTCGGCTTGCAGCAGATCTGCGAGCATGGGCGACCTTGGAGCAAGGCGGTGGCATC 656
DB 479 GCTCGGCTTGCAGCAGATCTGCGAGCATGGGCGACCTTGGAGCAAGGCGGTGGCATC 538
QY 657 CCGGCGGCGCCAGGCGGTGGGCGGCGGAGCTGTTGGAGGCGACCGGAGTGT 716
DB 539 CCGGCGGCGCCAGGCGGTGGGCGGCGGAGCTGTTGGAGGCGACCGGAGTGT 598
QY 717 GCGCTCAAGGAGCCTGGGCTCTGTGTGTTGTGATCATCAAGTTCTCAGAGCTCTCTACT 776
DB 599 GCGCTCAAGGAGCCTGGGCTCTGTGTGTTGTGATCATCAAGTTCTCAGAGCTCTCTACT 658
QY 777 GAGCGCTTGTGGGCGCACTACTGAGTGGCGCCCTGCTGCGAGGCC 822
DB 659 GAGCGCTTGTGGGCGCACTACTGAGTGGCGCCCTGCTGCGAGGCC 704

RESULT 7 634 bp mRNA linear EST 16-OCT-2001
BI907024 LOCUS 603064980F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214211 5',
DEFINITION mRNA sequence.
ACCESSION BI907024
VERSION BI907024.1 GI:16169804
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

Plate: L1AM1537 row: p column: 20
High quality sequence stop: 632.
Location/Qualifiers

FEATURES

source

1. 634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5214211"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC library."

ORIGIN

Query Match 64.0%; Score 582; DB 4; Length 634;
Best Local Similarity 100.0%; Pred. No. 1.2e-284;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GTGTCTTCAGACGCTATAGCTATGCACTCCAGCTCTTCAAGAGGAGGAGTAC 387
DB 32 GTGTCTTCAGACGCTATAGCTATGCACTCCAGCTCTTCAAGAGGAGGAGTAC 91
QY 388 TGCCTGTCGCGTGGGAGTCAAGCAGTCTGCAATTCACAGAGGCTCAGTGGAGACA 447
DB 92 TGCCTGTCGCGTGGGAGTCAAGCAGTCTGCAATTCACAGAGGCTCAGTGGAGACA 151
QY 448 GGTCTCCCGCCCAACCAAGCGGCGAGGCGGAGTCCGCGCGCCAGTGGTGGCCAGA 507
DB 152 GGTCTCCCGCCCAACCAAGCGGCGAGCGGCGGAGTCCGCGCGCCAGTGGTGGCCAGA 211
QY 508 CCGGCGGCGAGAGGCGGCCCGCAGCGCAGCCAGCAGCAGTCAAGCCCGCAGACTTCC 567
DB 212 CCGGCGGCGAGAGGCGGCCCGCAGCGCAGCCAGCAGCAGTCAAGCCCGCAGACTTCC 271
QY 568 TCTGAAGGCAAGTGAACCTTGAATCCGGCTCCGGTGGAGCAGAGTCTCGAGCAT 627
DB 272 TCTGAAGGCAAGTGAACCTTGAATCCGGCTCCGGTGGAGCAGAGTCTCGAGCAT 331
QY 628 GGGCGAGCTTGGAGCAGGCGGTGGCATCCGCGCGGCCCGCAGCGCTGCGCGGAGCTG 687
DB 332 GGGCGAGCTTGGAGCAGGCGGTGGCATCCGCGCGGCCCGCAGCGCTGCGCGGAGCTG 391
QY 688 GAGCTGTTGGGCGAGGCGCAGCGAGTGTGCTCAAGGAGCCTGGGCTGTGTGTTGT 747
DB 392 GAGCTGTTGGGCGAGGCGCAGCGAGTGTGCTCAAGGAGCCTGGGCTGTGTGTTGT 451
QY 748 GACATCAAGTTCTCAGAGCTCTCTACTGTGAAGCGCTTCTGGGCGGAGTCAAGTGGC 807
DB 452 GACATCAAGTTCTCAGAGCTCTCTACTGTGAAGCGCTTCTGGGCGGAGTCAAGTGGC 511
QY 808 GCCCTGTGAGGCGCCCGGCGGCGGTGTTCTGACTGAGGCGCTGCGAGAGGCTGTGGG 867
DB 512 GCCCTGTGAGGCGCCCGGCGGCGGTGTTCTGACTGAGGCGCTGCGAGAGGCTGTGGG 571
QY 868 CCGGAGGCTGTTGCGCTGCTGTCAGTGTGATGAGGCTGAC 909
DB 572 CCGGAGGCTGTTGCGCTGCTGTCAGTGTGATGAGGCTGAC 613

RESULT 8 905 bp mRNA linear EST 16-OCT-2001
BI910528 LOCUS 603068215F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217218 5',
DEFINITION mRNA sequence.
ACCESSION BI910528

```

VERSION      BI910528.1  GI:16173927
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 905)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cga@db-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://image.llnl.gov
              Plate: LLM11545 row: n column: 03
              High quality sequence stop: 863.

FEATURES
    source
        1..905
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5217218"
        /tissue_type="leukocyte"
        /lab_host="DH10B"
        /clone_1lb="NIH_MGC_118"
        /note="Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
        (destroyed); RNA source leukocytes from anonymous pool of
        non-activated adult donors. Library is oligo-dT primed
        and directionally cloned. (EcoRV site is destroyed upon
        cloning). Average insert size 1.7 kb. Insert size range
        1.2-3.3 kb. Library is normalized and enriched for
        full-length clones and was constructed by C. Gruber
        (Institute). Research Genetics tracking code 027. Note:
        this is a NIH_MGC Library."

ORIGIN
Query Match      64.0%; Score 582; DB 4; Length 905;
Best Local Similarity 100.0%; Pred. No. 1.2e-284;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GTGCTCTCCAGAACGTTAGCTATGAGCACTTCCAGCTTTTCAAGAGACAGAGGTTGCC 387
DB 38 GTGCTCTCCAGAACGTTAGCTATGAGCACTTCCAGCTTTTCAAGAGACAGAGGTTGCC 97
QY 388 TGCCGTGCGCGTGGCAGTCAAGCAGTTCTGCAATTCTCAGCAGGAGTCAAGTGGAGACA 447
DB 98 TGCCGTGCGCGTGGCAGTCAAGCAGTTCTGCAATTCTCAGCAGGAGTCAAGTGGAGACA 157
QY 448 GGCCTCCCCCAACCAAGCGGCGAGCGCGAGTCGGGCGCGCCCAAGTGTGTGCCAGA 507
DB 158 GGCCTCCCCCAACCAAGCGGCGAGCGCGAGTCGGGCGCGCCCAAGTGTGTGCCAGA 217
QY 508 CGGCGCGGAGAGGGGCGCCGACCCGACCCAGCAGCAGTCAAGCCCGCCAGACTTCC 567
DB 218 CGGCGCGGAGAGGGGCGCCGACCCGACCCAGCAGCAGTCAAGCCCGCCAGACTTCC 277
QY 568 TCTGAAGGCAAGTACCTGTGACATCGGGCTCCGGGTTTCAGAGCAGTACTCGAGAGAT 627
DB 278 TCTGAAGGCAAGTACCTGTGACATCGGGCTCCGGGTTTCAGAGCAGTACTCGAGAGAT 337
QY 628 GGGCCAGCTTTGAGCAGAGGCGTGGCATCCGCGCGCCCAAGCGCTGCGCGAGAGCTG 687
DB 338 GGGCCAGCTTTGAGCAGAGGCGTGGCATCCGCGCGCCCAAGCGCTGCGCGAGAGCTG 397
QY 688 GACGTGTTTGGGCAAGGCCACCGCAGTGTGGGCTCAAGGACCTTGGGCTTGTGTGTGT 747
DB 398 GACGTGTTTGGGCAAGGCCACCGCAGTGTGGGCTCAAGGACCTTGGGCTTGTGTGTGT 457

```

```

QY 748 GACATCAAGTTCAGAGCTCTCTATCTGAGACGCTTTCGGGGGCACTACCTGAGTGGC 807
DB 458 GACATCAAGTTCAGAGCTCTCTATCTGAGACGCTTTCGGGGGCACTACCTGAGTGGC 517
QY 808 GCCCTGCTGACAGGCGCTTCGGGGGCGTGTCTTCTACTAGAGCCCTTCGAGAGGCTGTGGGC 867
DB 518 GCCCTGCTGACAGGCGCTTCGGGGGCGTGTCTTCTACTAGAGGCGCTTCGAGAGGCTGTGGGC 577
QY 868 CGGAGGCTGTTCCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
DB 578 CGGAGGCTGTTCCGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 619

RESULT 9
LOCUS      BG251151
DEFINITION BG251151 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473349 5',
            mRNA sequence.
ACCESSION  BG251151
VERSION     BG251151.1  GI:12760967
KEYWORDS
SOURCE      EST.
ORGANISM    Homo sapiens (human)
REFERENCE    1 (bases 1 to 927)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cga@db-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNLN at:
              http://image.llnl.gov
              Plate: LLM10295 row: k column: 14
              High quality sequence stop: 715.

FEATURES
    source
        1..927
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4473349"
        /tissue_type="adenocarcinoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /clone_1lb="NIH_MGC_90"
        /note="Organ: liver; Vector: PCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
        Average insert size 1.7 kb. Library enriched for
        full-length clones and constructed by Life Technologies.
        Note: this is a NIH_MGC Library."

ORIGIN
Query Match      64.0%; Score 582; DB 4; Length 927;
Best Local Similarity 100.0%; Pred. No. 1.2e-284;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GTGCTCTCCAGAACGTTAGCTATGAGCACTTCCAGCTTTTCAAGAGACAGAGGTTGCC 387
DB 60 GTGCTCTCCAGAACGTTAGCTATGAGCACTTCCAGCTTTTCAAGAGACAGAGGTTGCC 119
QY 388 TGCCGTGCGCGTGGCAGTCAAGCAGTTCTGCAATTCTCAGCAGGAGTCAAGTGGAGACA 447
DB 120 TGCCGTGCGCGTGGCAGTCAAGCAGTTCTGCAATTCTCAGCAGGAGTCAAGTGGAGACA 179
QY 448 GGCCTCCCCCAACCAAGCGGCGAGCGGAGTGTGGGCGCGCCCAAGTGTGTGTGCCAGA 507
DB 180 GGCCTCCCCCAACCAAGCGGCGAGCGGAGTGTGGGCGCGCCCAAGTGTGTGTGCCAGA 239
QY 508 CGGCGCGGAGAGGGGCGCCGACCCGACCCAGCAGCAGTCAAGAGCCCGCCAGACTTCC 567

```

|||||
240 CCGGGGCGAGAGGGGCCCGACCGACCGACGAGCTGAGAGCCCGACACTTCC 299
QY 568 TCTGAAGCAAAAGTACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGGACAT 627
Db 300 TCTGAAGCAAAAGTACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGGACAT 359
QY 628 GGGCGACCTTTGAGAGAGGGGCTGGCATCCCGCGGCCCGACAGGCGCTGGCGGCACTG 687
Db 360 GGGCGACCTTTGAGAGAGGGGCTGGCATCCCGCGGCCCGACAGGCGCTGGCGGCACTG 419
QY 688 GAGCTGTTGGGCGAGGCGACCGAGTGTGGCTCAAGGAGCACTGGGCTGTGGTTGT 747
Db 420 GACGTGTTGGGCGAGGCGACCGAGTGTGGCTCAAGGAGCACTGGGCTGTGGTTGT 479
QY 748 GACATCAAGTTCTGAGAGCTCTCTATCTGAGAGCTTGGGCGGAGTACTGAGTGGC 807
Db 480 GACATCAAGTTCTGAGAGCTCTCTATCTGAGAGCTTGGGCGGAGTACTGAGTGGC 539
QY 808 GCCCTGCTGAGAGGCTTGGCGGGCGTGTCTTCTGACTGAGGCTTGGAGAGGCTGTGGC 867
Db 540 GCCCTGCTGAGAGGCTTGGCGGGCGTGTCTTCTGACTGAGGCTTGGAGAGGCTGTGGC 599
QY 868 CGGAGGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
Db 600 CGGAGGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 641

RESULT 10
BI910416 947 bp mRNA linear EST 16-OCT-2001
LOCUS 603068087P1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217213 5',
DEFINITION mRNA sequence.
ACCESSION BI910416
VERSION BI910416.1 GI:16173804
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 947)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM1545 row: m column: 22
High quality sequence stop: 853.
Location/Qualifiers
1. 947
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5217213"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/clone_id="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source: leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 64.0%; Score 582; DB 4; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.2e-284;
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 328 GTGCTTCACAAAGGCTATAGCTATAGGACCTCCAGCTCTTCAAGAGGAGCAGAGGGTAC 387
Db 72 GTGCTTCACAAAGGCTATAGCTATAGGACCTCCAGCTCTTCAAGAGGAGAGGTTAGC 131
QY 388 TGCGTTCGCGGTGGGCGAGTCAAGCAGTTCTGCAATTTCTGAGAGGGTCAAGTGGAGACA 447
Db 132 TGCGTTCGCGGTGGGCGAGTCAAGCAGTTCTGCAATTTCTGAGAGGGTCAAGTGGAGACA 191
QY 448 GGTCTCCCCCAACCAAGCGGCGAGCGGAGTGGGGCGCGGCCAGGTGTGCCAGA 507
Db 192 GGTCTCCCCCAACCAAGCGGCGAGCGGAGTGGGGCGCGGCCAGGTGTGTGCCAGA 251
QY 508 CGGCGGCGAGAGGGGCGCGGCGGACCGGACAGAGTCAAGAGCCCGGACAGCTTCC 567
Db 252 CGGCGGCGAGAGGGGCGCGGCGGACCGGACAGAGTCAAGAGCCCGGACAGCTTCC 311
QY 568 TCTGAAGCAAAAGTACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGAGAT 627
Db 312 TCTGAAGCAAAAGTACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGAGAT 371
QY 628 GGGCGACCTTTGAGAGAGGCGTGGCATCCCGCGGCCCGACGCGCTGGCGGCGAGCTG 687
Db 372 GGGCGACCTTTGAGAGAGGCGTGGCATCCCGCGGCCCGACGCGCTGGCGGCGAGCTG 431
QY 688 GAGCTGTTGGGCGAGGCGACCGAGTGTGGCTCAAGGAGCACTGGGCTGTGGTTGT 747
Db 432 GAGCTGTTGGGCGAGGCGACCGAGTGTGGCTCAAGGAGCACTGGGCTGTGGTTGT 491
QY 748 GACATCAAGTTCTGAGAGCTCTCTATCTGAGAGCTTGGGCGGAGTACTGAGTGGC 807
Db 492 GACATCAAGTTCTGAGAGCTCTCTATCTGAGAGCTTGGGCGGAGTACTGAGTGGC 551
QY 808 GCCCTGCTGAGAGGCTTGGCGGGCGTGTCTTCTGACTGAGGCTTGGAGAGGCTGTGGC 867
Db 552 GCCCTGCTGAGAGGCTTGGCGGGCGTGTCTTCTGACTGAGGCTTGGAGAGGCTGTGGC 611
QY 868 CGGAGGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909
Db 612 CGGAGGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 653

RESULT 11
BM908148 1062 bp mRNA linear EST 12-MAR-2002
LOCUS AGENCOURT_6707554 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5745179
DEFINITION 5', mRNA sequence.
ACCESSION BM908148
VERSION BM908148.1 GI:19358527
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1062)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12767 row: 1 column: 12

Db	510	GCCCTGCTGCAGGCCCTTCGGGGCGCTGTCTTCTGACTGACGGCCCTTGCCGAGAGCCTGTGTGGC	569
OY	868	CGGAGGCTGTTCGCTGCTGCTGATCATGTTGATGATGAGCTGAC	909
Db	570	CGGAGGCTGTTCGCTGCTGCTGATGATGATGAGCTGAC	611
RESULT_13			
CRS99685			
LOCUS			
DEFINITION	full-length cDNA clone CSOD1069YH17 of Placenta Cot 25-normalized		
ACCESSION	CRS99685	1510 bp	mRNA linear HTC 21-JUL-2004
VERSION	CRS99685.1		
KEYWORDS	HTC; CNSLT cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
JOURNAL	Full-length cDNA libraries and normalization		
REMARK	Unpublished		
	Contact : Peng Liang Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
REFERENCE	Faraday Avenue		
AUTHORS	2 (bases 1 to 1510)		
TITLE	Genoscope.		
JOURNAL	Direct Submission		
	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :		
	Bp 191 91006 EVRI cedex - FRANCE (E-mail : seque@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)		
	1st strand cDNA was primed with a NotI-oligo(dt) primer. Five primers		
	end enriched, double-strand cDNA was digested with Not I and cloned		
	into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library		
	was normalized. Library was constructed by life technologies, a		
	division of invitrogen.		
FEATURES	Location/Qualifiers		
source	1..1510		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSOD1069YH17"		
	/tissue_type="Placenta Cot 25-normalized"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
Query Match	64.0%; Score 582; DB 3; Length 1510;		
Best Local Similarity	100.0%; Pred. No. 1.2e-284; Indels 0; Gaps 0;		
Matches	582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	328	GTGTTCAGAAAGCTATAGTATGACCTTCACAGCTCTTCAAGAAGAGAGGGTAGC	387
Ddb	29	GTGTTCAGAAAGCTATAGTATGACCTTCACAGAGAGGGTAGC	88
OY	388	TGCGGTGCGCGTCGGCAGTCAAAGATTCTGCAATTCTCAGCAGAGGTCAAGTGGAGACA	447
Ddb	89	TGCGGTGCGCGTCGGCAGTCAAAGATTCTGCAATTCTCAGCAGAGGTCAAGTGGAGACA	148
OY	448	GGCTTCCCCCAACCAAGCGGAGCGGCGAAGTCGGGGCCGAGTGTGTGTCCACA	507
Ddb	149	GGCTTCCCCCAACCAAGCGGAGCGGCGAAGTCGGGGCCGAGTGTGTGTCCACA	208
OY	508	CGGCGGCGGAAGGGGGCCCGAGCGCACCCAGCAGCAGTAAGAGCCGCGCAGACCTTCC	567
Ddb	209	CGGCGGCGGAAGGGGGCCCGAGCGCACCCAGCAGCAGTAAGAGCCGCGCAGACCTTCC	268
OY	568	TCTGAAGCAAAAGTACCTGTGTACATCCGGCTCCGGGTTTCGAGCAGATGACTGCAGCAT	627
Ddb	269	TCTGAAGCAAAAGTACCTGTGTACATCCGGCTCCGGGTTTCGAGCAGATGACTGCAGCAT	328
OY	628	GGGCGAGCCTTGAGCAGGGGCTGGGCATTCGGCGGGCCCAAGCGCTGCGCGGACGCTG	687

Df	329	GggccAGcCTTGGAgCAGgGCGGTGSCATrCCCGGCGGCCCAAGAgCGCTGGCGCGGCACTG	388
OY	688	GACCTGTTTGgGCAGGcCACCGCAGTCGTGCGCTCAAGAgGACCTGGcCTGTGTGTGT	747
Df	389	GAGCGTATTGGgGCAGGcCACCGCAGTCGTGCGCTCAAGAgGACCTGGcCTGTGTGTGT	448
OY	748	GACATCAAGTTCTAGAGCTTCCTTAfCTGGAAGcCCTTCCTGGGgGCATACCTGAGfTGGC	807
Df	449	GACATCAAGTTCTAGAGCTTCCTTAfCTGGAAGcCCTTCCTGGGgGCATACCTGAGfTGGC	508
OY	808	GGCGTGCAGGAGCGCCCTGGCGGgCGGTGTCCTGACTAGAGcCCTGGAGAgGCTGTGGG	867
Df	509	GCCCTGTGAGGCGcCCTGGCGGgCGGTGTCCTGACTAGAGcCCTGGAGAgGCTGTGGG	568
OY	868	CGGAGGCTGTTCGCTGCTGCTGTCAGTGTGATGAGGCTGAC	909
Df	569	CGGAGGCTGTTCGCTGCTGCTGTCAGTGTGATGAGGCTGAC	610
RESULT 14			
LOCUS	CR590827	1586 bp	mRNA linear HTC 21-JUL-2004
DEFINITION	Full-length cDNA clone CSODBD01W11 of Placenta of Homo sapiens		
ACCESSION	(human).		
VERSION	CR590827.1 GI:50471634		
KEYWORDS	HTC; CNSLT cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
REMARK	Contact : Feng Liang Email : filiang@lifeitech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1586) Genoscope.		
FEATURES	Direct Submission		
source	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a Notti-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		
ORIGIN	Location/Qualifiers 1..1586 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSODBD01W11" /tissue_type="Placenta" /plasmid="pCMVSPORT_6"		
Query Match	64.0%; Score 582; DB 3; Length 1586;		
Best Local Similarity	100.0%; Prod. No. 1.2e-284;		
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	328	GMGTCCTCAGAAGCGTATAGCTATAGGACCCTCCAGCTCTTCAAAGAGGAGACAGGGTATGC	387
Df	111	GTGTCTCAGAACCGTATAGCTATAGGACCCTCCAGCTCTTCAAAGAGGAGACAGGGTATGC	170
OY	388	TGCCGTGCCCTCGGcAGTCAAGAGATTCTTCAATTCTCAGCAGGGTCACTGTGGAGACA	447
Df	171	TGCCGTGCCCTCGGcAGTCAAGAGATTCTTCAATTCTCAGCAGGGTCACTGTGGAGACA	230
OY	448	GAGTCCTCCCCCAACCAAGCGGcAGCGGAGTGTGGGGCGGCCCAAGTGTGTGCCAGA	507

Db 231 GGCTCCCCCAACCAAGGCGAGCGGAGTGGGGCCGAGCCCACTGTTGTGCCAGA 290
 Qy 508 CGGCGCGAGAGAGGGGCCCGCAGCCGACCCGAGAGAGAGCCCGCAGACTTCC 567
 Db 291 CGGCGCGAGAGAGGGGCCCGCAGCCGACCCGAGAGAGAGCCCGCAGACTTCC 350
 Qy 568 TCTGAAGGCAAGTGAAGTCTGTGACATCCGGCTCCGGGTTGAGAGAGAGTACTGAGAGAT 627
 Db 351 TCTGAAGGCAAGTGAAGTCTGTGACATCCGGCTCCGGGTTGAGAGAGAGTACTGAGAGAT 410
 Qy 628 GGGGCGAGCTTGAAGAGAGGCGTGGACATCCGGGGGGCCCGAGGGCTGGCGCGAGAGTGG 687
 Db 411 GGGGCGAGCTTGAAGAGAGGCGTGGACATCCGGGGGGCCCGAGGGCTGGCGCGAGAGTGG 470
 Qy 688 GAGCTGTTTGGGCGAGGCGACCGCAGTGTGCTCAAGAGGACTGGGCTGTGTGTGTGT 747
 Db 471 GAGCTGTTTGGGCGAGGCGACCGCAGTGTGCTCAAGAGGACTGGGCTGTGTGTGTGTGT 530
 Qy 748 GACTCAAGTCTCAAGAGCTTCTCTATCTGGAGCGCTTCTGGGCGAGTACTGAGTGGC 807
 Db 531 GACTCAAGTCTCAAGAGCTTCTCTATCTGGAGCGCTTCTGGGCGAGTACTGAGTGGC 590
 Qy 808 GCGCTGTGAGAGGCGCTGGGGGGGGTCTCTGACTGAGGCGCTGGAGAGGCTGTGGGC 867
 Db 591 GCGCTGTGAGAGGCGCTGGGGGGGGTCTCTGACTGAGGCGCTGGAGAGGCTGTGGGC 650
 Qy 868 CGGAGAGCTGTTCCGCTGCTGTGTGAGTGTGATGAGGCTGAC 909
 Db 651 CGGAGAGCTGTTCCGCTGCTGTGTGAGTGTGATGAGGCTGAC 692

RESULT 15
 LOCUS BM925969 1089 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT 6649774 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764272
 5', mRNA sequence.
 ACCESSION BM925969
 VERSION BM925969.1 GI:19376336
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LLM12817 row: h column: 01
 High quality sequence start: 14
 High quality sequence stop: 586.
 Location/Qualifiers
 1..1089

FEATURES
 Source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5764272"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_114"
 /note="Organ: brain; Vector: pCMV-SportE; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. library is normalized and enriched for

full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC library."

Query Match 61.4%; Score 558; DB 5; Length 1089;
 Best Local Similarity 99.7%; Pred. No. 2e-272;
 Matches 658; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN
 1 ATGGCGCTATCCGGGTGAGACCCCGGCGGCTGGGAGAGTATGCTGAGTAC 60
 71 ATGGCGCTATCCGGGTGAGACCCCGGCGGCTGGGAGAGTATGCTGAGTAC 130
 61 TACGGATGCTGCTGCTTCAACCGTATGTTGAGGTGTGGCGGAGCACTGACGAGTGC 120
 131 TACGGATGCTGCTGCTTCAACCGTATGTTGAGGTGTGGCGGAGCACTGACGAGTGC 190
 121 GAGCTGAGGCTCTGGCTTTCTGCTGATGAGGCTCTGGCGCGCGGAGGCTTACCC 180
 191 GAGCTGAGGCTCTGGCTTTCTGCTGATGAGGCTCTGGCGCGCGGAGGCTTACCC 250
 181 CGGGCGGCGAGCGGCTGAGAGCTCTGCTGAGAGCTGAGAGCGCGCGGAGGCTGAG 240
 251 CGGGCGGCGAGCGGCTGAGAGCTCTGCTGAGAGCTGAGAGCGCGCGGAGGCTGAG 310
 241 AGCAACCTGCGGCTGCTGGGCAACTCTGCGCGCTGCGCCGCAAGACTCTGCGCG 300
 311 AGCAACCTGCGGCTGCTGGGCAACTCTGCGCGCTGCGCCGCAAGACTCTGCGCG 370
 301 CACTGCGCGGCAAGCGCGCGCGGCAAGTCTTCAAGAGCTATGAGCACTTC 360
 371 CACTGCGCGGCAAGCGCGCGCGGCAAGTCTTCAAGAGCTATGAGCACTTC 430
 361 AGCTCTCAAGAGAGAGAGAGAGTGTGCTGCGCGCTGCGCGAGTCAAGAGTTCGCA 420
 431 AGCTCTCAAGAGAGAGAGAGTGTGCTGCGCGCTGCGCGAGTCAAGAGTTCGCA 490
 421 AATTCTCAGAGAGGTCAGTGGAGAGAGAGCTCCCGCCCAACCAAGCGGAGCGGAGT 480
 491 AATTCTCAGAGAGGTCAGTGGAGAGAGAGCTCCCGCCCAACCAAGCGGAGCGGAGT 550
 481 CGGGCGGCGGCGGAGTGTGCTGCGAGAGCGGCGGAGAGAGGCGCGGAGCGGAG 540
 551 CGGGCGGCGGCGGAGTGTGCTGCGAGAGCGGCGGAGAGAGGCGCGGAGCGGAG 610
 541 CAGAGTCAGAGAGCGGCGGAGAGCTTCTTGAAGGCAAGTGAATCTGATCCGCTC 600
 611 CAGAGTCAGAGAGCGGCGGAGAGCTTCTTGAAGGCAAGTGAATCTGATCCGCTC 670
 601 CGGTTTGAAGAGAGTACTGCGAGCATGGGCGGAGCTTGGAGAGGCGGAGTCCCGG 660
 671 CGGTTTGAAGAGAGTACTGCGAGCATGGGCGGAGCTTGGAGAGGCGGAGTCCCGG 730

RESULT 16
 LOCUS CK429257 655 bp mRNA linear EST 06-JAN-2004
 DEFINITION o334b12.y1 Human lacrimal gland, unamplified: o3 Homo sapiens cDNA
 clone o334b12 5', mRNA sequence.
 ACCESSION CK429257
 VERSION CK429257.1 GI:40676160
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Eye Institute
 Unpublished (2002)
 Contact: Wiscow G
 Section on Molecular Structure and Function

6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: Graeme@helix.nih.gov
 Plate: 34 row: b column: 12
 Seq primer: M13RPL reverse primer (ABT).
 Location/Qualifiers

FEATURES

source

1..655
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="oj34b12"
 /tissue_type="lacrimal gland"
 /dev_stage="Adult"
 /lab_host="EMD10B"
 /clone_1ib="Human lacrimal gland, unamplified: oj"
 /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted from 2 human lacrimal glands. A directionally cloned cDNA library in the pCMVSPORT6 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGGCGGCCGCT(7)15-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 58.5%; Score 532; DB 7; Length 655;
 Best Local Similarity 99.8%; Pred. No. 3.4e-259;
 Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCGGTATCCGGGTGACCCCGGCGTGGAGGAGATGAGTGGCTGATAC 60
 73 ATGGCGGTATCCGGGTGACCCCGGCGTGGAGGAGATGAGTGGCTGATAC 132
 61 TACGGGATGCTGCTTCAACCTATGTTGAGGTGTTGGCGGCAATGACCGAGTGC 120
 133 TACGGGATGCTGCTTCAACCTATGTTGAGGTGTTGGCGGCAATGACCGAGTGC 192
 121 GAGCTGAGTCTCTGCGCTTTCTGCTGAGTGAAGCTCTCTGCGCGCGGAGCTTGGCC 180
 193 GAGCTGAGTCTCTGCGCTTTCTGCTGAGTGAAGCTCTCTGCGCGCGGAGCTTGGCC 252
 181 CGGGCCCGGAGCGGCGCTAGAGCTCTGCTGAGTGAAGCGCGGCGAGTGGCGGAG 240
 253 CGGGCCCGGAGCGGCGCTAGAGCTCTGCTGAGTGAAGCGCGGCGAGTGGCGGAG 312
 241 AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCTGCTGGCCGCAAGACTGCTCCG 300
 313 AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCTGCTGGCCGCAAGACTGCTCCG 372
 301 CACTGGCGGCAAGCGGCGCGGCGCACTGCTCTCAAGAGCTATAGTATGCACTCC 360
 373 CACTGGCGGCAAGCGGCGCGGCGCACTGCTCTCAAGAGCTATAGTATGCACTCC 432
 361 AGCTCTCAAGAGAGAGAGGTAAGTCCGCTGCGGCGGCGAGTGAAGAGTTTGA 420
 433 AGCTCTCAAGAGAGAGAGGTAAGTCCGCTGCGGCGGCGAGTGAAGAGTTTGA 492
 421 AATTCTCAGCAGGCTCAGTGGAGAGAGCTCCCCCAACCAAGCGGAGCGGAGT 480
 493 AATTCTCAGCAGGCTCAGTGGAGAGAGCTCCCCCAACCAAGCGGAGCGGAGT 552
 481 CGGGCGGCGGCGCACTGCTGCTGCAAGCGGCGGAGAGGAGGCGCCAGCGCAG 540
 553 CGGGCGGCGGCGCACTGCTGCTGCAAGCGGCGGAGAGGAGGCGCCAGCGCAG 612
 541 CAGAGTCAGAGCGCGCAGACTTCTCTGAAGGCAAGTGA 583
 613 CAGAGTCAGAGCGCGCAGACTTCTCTGAAGGCAAGTGA 655

RESULT 17
 CN373290 644 bp mRNA linear EST 16-MAY-2004
 LOCUS 17000424182765 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
 DEFINITION CN373290
 ACCESSION CN373290.1 GI:47373224
 VERSION EST.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Li, Y., Xu, C., Fang, R., Guejler, K., Rao, M. S., Mandalam, R., Lebkowski, J and Stanton, L. W.
 Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 CONTACT: Brandenberger R
 Regenerative Medicine
 Genon Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@genon.com
 Insert length: 644 Std Error: 0.00.
 Location/Qualifiers

1..644
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
 /note="Toigo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

FEATURES

source

Query Match 58.4%; Score 531; DB 7; Length 644;
 Best Local Similarity 99.8%; Pred. No. 1.1e-258;
 Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

328 GTGTCTCCAGAGAGCTATAGCTATGACCTCCAGCTTTCAAGAGAGAGAGTAC 387
 30 GTGTCTCCAGAGAGCTATAGCTATGACCTCCAGCTTTCAAGAGAGAGAGTAC 89
 388 TGGCGTGGCGGCTGCGGAGTCAAGAGTTTCTGCAATTCTCAGCAGGCTCAGTGGAGACA 447
 90 TGGCGTGGCGGCTGCGGAGTCAAGAGTTTCTGCAATTCTCAGCAGGCTCAGTGGAGACA 149
 448 GGGTCCCCCGCAACCAAGCGGCGGAGGAGTGGGGCGGCGGCGGAGTGGTGGAGACA 507
 150 GGGTCCCCCGCAACCAAGCGGCGGAGGAGTGGGGCGGCGGCGGAGTGGTGGAGACA 209
 508 CGGGCGGAGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 567
 210 CGGGCGGAGAGAGGCGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 269
 270 TCTGAAGGCAAGAGTGAATCTGATCGGCTCGGGTTCGAGCAGAGTACTGGAACAT 329
 568 TCTGAAGGCAAGAGTGAATCTGATCGGCTCGGGTTCGAGCAGAGTACTGGAACAT 627
 628 GGGCGAGCTTTGAGAGAGGCGGCGGAGTCCCGGCGGCGGCGGAGGCGGCGGCGGAGCTG 687
 330 GGGCGAGCTTTGAGAGAGGCGGCGGAGTCCCGGCGGCGGCGGAGGCGGCGGCGGAGCTG 389
 688 GAGCTGTTTGGAGAGGCGGAGAGTCTGCTCAAGGAGAGTCTGTTGTTGT 747
 390 GAGCTGTTTGGAGAGGCGGAGAGTCTGCTCAAGGAGAGTCTGTTGTTGT 449

Qy 748 GACATCAAGTTCTAGAGCTTCTCTATCTGAGACGCTTCTGGGGGAGCTACCTAGTGGC 807
|
|
|
Db 450 GACATCAAGTTCTAGAGCTTCTCTATCTGAGACGCTTCTGGGGGAGCTACCTAGTGGC 509
|
|
|
Qy 808 GCCCTGTCGAGCGCCCTGGGGGGGTGTTCTCTGACTGAGGCGCTCGAGAGGCTGTGGC 867
|
|
|
Db 510 GCCCTGTCGAGCGCCCTGGGGGGGTGTTCTCTGACTGAGGCGCTCGAGAGGCTGTGGC 569
|
|
|
Qy 868 CGGAGGCTGTTCGCTGCTGTGATGATGATGAGGCTGAC 909
|
|
|
Db 570 CGGAGGCTGTTCGCTGCTGTGATGATGATGAGGCTGAC 611
|
|
|
RESULT 18
BM559141
LOCUS BM559141 993 bp mRNA linear EST 20-FEB-2002
DEFINITION AGENCOURT_6562852 NIH_MGC_67 Homo sapiens cDNA IMAGE:5551258
5', mRNA sequence.
ACCESSION BM559141
VERSION BM559141.1 GI:18802480
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 993)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LHAM12264 row: 1 column: 11
High quality sequence start: 31
High quality sequence stop: 611.
Location/Qualifiers
1. .993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5551258"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 Kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 58.4%; Score 531; DB 4; Length 993;
Best Local Similarity 100.0%; Pred. No. 1.1e-258;
Matches 531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 379 GAGGATGATGCGCTGCGCGCTGCGAGTCAAGCAATTCGACAGGCTGAC 438
|
|
|
Db 41 GAGGATGATGCGCTGCGCGCTGCGAGTCAAGCAATTCGACAGGCTGAC 100
|
|
|
Qy 439 TGGGAGACAGGCTCCCCCAACCAAGGCGGCGGAGTGGGGCCGCGCCCAATGCT 498
|
|
|
Db 101 TGGGAGACAGGCTCCCCCAACCAAGGCGGCGGAGTGGGGCCGCGCCCAATGCT 160
|
|
|
Qy 499 GGTGCGAGCGCGCGGAGAGGGGCGCCGACCGACCCGACGAGTCAAGGCGCGCC 558
|
|
|
Db 161 GGTGCGAGCGCGCGGAGAGGGGCGCCGACCGACCCGACGAGTCAAGGCGCGCC 220
|
|
|
Qy 559 AGACCTTCTCTGAAGCGAAGTGAATCCGGCTCCGGGTTTCAGAGAGTAC 618
|
|
|

Db 221 AGACCTTCTCTGAAGCGAAGTGAATCCGGCTCCGGGTTTCAGAGAGTAC 280
|
|
|
Qy 619 TGGGAGATGGGCGAGGCTTTGAGACAGGCGGTGATCCGGCGGCGCCCAAGCGCTGGC 678
|
|
|
Db 281 TGGGAGATGGGCGAGGCTTTGAGACAGGCGGTGATCCGGCGGCGCCCAAGCGCTGGC 340
|
|
|
Qy 679 CGGAGCTGAGCATGTTTGGGCGAGCGCACCGAGTGTGCTGCTCAAGGAGCCTGGGCTCT 738
|
|
|
Db 341 CGGAGCTGAGCATGTTTGGGCGAGCGCACCGAGTGTGCTGCTCAAGGAGCCTGGGCTCT 400
|
|
|
Qy 739 GTGGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTGTGGGGGACTAC 798
|
|
|
Db 401 GTGGTTGTGACATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTGTGGGGGACTAC 460
|
|
|
Qy 799 CTGAGTGGCGCCCTGTCGACAGGCGCCCTGCGGGGCGGTTCCTGACGAGGCGCTGGAG 858
|
|
|
Db 461 CTGAGTGGCGCCCTGTCGACAGGCGCCCTGCGGGGCGGTTCCTGACGAGGCGCTGGAG 520
|
|
|
Qy 859 GCTGTGGGCGCGAGGCTGTTCGCTGCTGATGATGAGGCTGAC 909
|
|
|
Db 521 GCTGTGGGCGCGAGGCTGTTCGCTGCTGATGATGAGGCTGAC 571
|
|
|
RESULT 19
B0073197
LOCUS B0073197 1051 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6761557 NIH_MGC_120 Homo sapiens cDNA IMAGE:5746449
5', mRNA sequence.
ACCESSION B0073197
VERSION B0073197.1 GI:19902243
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1051)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LHAM12771 row: a column: 10
High quality sequence stop: 651.
Location/Qualifiers
1. .1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5746449"
/lab_host="DH10B"
/clone_id="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 Kb, insert size range 1-2.5 Kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 58.4%; Score 531; DB 5; Length 1051;
Best Local Similarity 99.8%; Pred. No. 1.1e-258;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	328	GTGCTCTCAGAAACGCTTAGCTATAGGGACACCTCCAGCTCTTCAAGAGAGACAAGAGGGTAGC	387
Db	62	GTGCTCTCAGAAACGCTTAGCTATAGGGACACCTCCAGCTCTTCAAGAGAGACAAGAGGGTAGC	121
QY	388	TGCCGTGCGCGTCGGCAGCTCAAGCAGATTCTTGCAAAATTTCTAGCAGAGGGTCAGTGGAGACA	447
Db	122	TGCCGTGCGCGTCGGCAGCTCAAGCAGATTCTTGCAAAATTTCTAGCAGAGGGTCAGTGGAGACA	181
QY	448	GGCTTCCCCCCCCAACCAAGCGGCACAGCGCGAGATGGGGGCGCGCCCACTGTGGTGGCCAGA	507
Db	182	GGCTTCCCCCCCCAACCAAGCGGCACAGCGCGAGATGGGGGCGCGCCCACTGTGGTGGCCAGA	241
QY	508	CGGCGGCGAGAGAGGGGCGCCAGCGCGCAACCCAGAGAGAGTGAAGACCOCGACAGCTTTC	567
Db	242	CGGCGGCGAGAGAGGGGCGCCAGCGCGCAACCCAGAGAGAGTGAAGACCOCGACAGCTTTC	301
QY	568	TTGTGAAGCGAAAGTGACCTGTGACATTCGAGCTCGGCTTCGGGTTTGAGAGAGTACTGGAGCAT	627
Db	302	TTGTGAAGCGAAAGTGACCTGTGACATTCGAGCTCGGCTTCGGGTTTGAGAGAGTACTGGAGCAT	361
QY	628	GGGCGACACCTTTGAGAGCAGGGGTGGCATCCCGGCGGCCCAAGGGGCGTGGCGCGGACACTG	687
Db	362	GGGCGACACCTTTGAGAGCAGGGGTGGCATCCCGGCGGCCCAAGGGGCGTGGCGCGGACACTG	421
QY	688	GACGTGTTTGGGCGAGGCCACCGCAGTCTGCGCTCAAGAGGACTGTGGGCTCTGTGTTTGT	747
Db	422	GACGTGTTTGGGCGAGGCCACCGCAGTCTGCGCTCAAGAGGACTGTGGGCTCTGTGTTTGT	481
QY	748	GACATCAAGTTCTCAGAGCTCTCCTATCTGAGCGCTTTCTGGGGCGACTACTAGTGGC	807
Db	482	GACATCAAGTTCTCAGAGCTCTCCTATCTGAGCGCTTTCTGGGGCGACTACTAGTGGC	541
QY	808	GCCCTGCTGACAGGCGCCCTGGGGGGGTGTTCCTGACTAGAGGCCCTGGCAGAGGCTGTGGGC	867
Db	542	GCCCTGCTGACAGGCGCCCTGGGGGGGTGTTCCTGACTAGAGGCCCTGGCAGAGGCTGTGGGC	601
QY	868	CGGAGAGGCTGTTCCGCTGCTGGTGTCAGTGTGATGAGGCTGAC	909
Db	602	CGGAGAGGCTGTTCCGCTGCTGGTGTCAGTGTGATGAGGCTGAC	643
RESULT 20			
LOCUS	BU902131	1161 bp	mRNA linear EST 17-OCT-2002
DEFINITION	AGENCOURT 10127738 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502805		
ACCESSION	BU902131		
VERSION	BU902131.1	GI:24084044	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1161)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LIML at:		
	http://image.liml.gov		
	plate: L1AM14059 row: h column: 06		
	High quality sequence stop: 707.		
FEATURES			
SOURCE	location/Qualifiers		
	1..1161		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		

```

/db xref="taxon:9606"
/clone="IMAGE:6502805"
/tissue_type="leiomyosarcoma"
/lab_host="RD10B (phage-resistant)"
/clone_1jb="NTH MGC 71"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Size_1: Not1;
Site_2: Salt; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

```

	Query Match	Best Local Similarity	58.4%; Score 531; DB 5; Length 1161;	99.8%; Pred. No. 1,1e-258;	Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	328	GTGTCTTCAGAACGGCTATAGCTATGSCACTTCCAGCTCTTCCAAAGAGACAGAGGGTAGC	387		
Db	77	GTGTCTTCAGAACGGCTATAGCTATGSCACTTCCAGCTCTTCCAAAGAGAGAGAGGGTAGC	136		
Qy	388	TGCGCTGCGCGTGGGCGAGTCAAGAGATTCTGCAATTCTCAGACGAGGTTCAGTGGGAACA	447		
Db	137	TGCGCTGCGCGTGGGCGAGTCAAGAGATTCTGCAATTCTCAGACGAGGTTCAGTGGGAACA	196		
Qy	448	GGCTCCCCCCCCCAACCAAGCGGACAGCGCGGAGTCCGAGGCGCGGCCCACTGGTGTGCCAGA	507		
Db	197	GGCTCCCCCCCCCAACCAAGCGGACAGCGCGGAGTCCGAGGCGCGGCCCACTGGTGTGCCAGA	256		
Qy	508	CGGCGGCGAGAGAGGGGCCCGCAGCCGACCCCAAGACGAGTCAAGACCCGCGACACTTCC	567		
Db	257	CGGCGGCGAGAGAGGGGCCCGCAGCCGACCCCAAGACGAGTCAAGACCCGCGACACTTCC	316		
Qy	568	TCTGAAGGCAAAAGTGAACCTGTGACATCCGGCTCCGGTTTGAGACAGATCTCGGACAT	627		
Db	317	TCTGAAGGCAAAAGTGAACCTGTGACATCCGGCTCCGGTTTGAGACAGATCTCGGACAT	376		
Qy	628	GGGCGACGCTTTGAGAGCGGCGGTGACATCCCGGCGGCGCCAGGCGCTGCGCGGCACTG	687		
Db	377	GGGCGACGCTTTGAGAGCGGCGGTGACATCCCGGCGGCGCCAGGCGCTGCGCGGCACTG	436		
Qy	688	GACGTGTTTGGGCGAGGCGCACCGACGTGCTGCGCTCAAGAGGACCTGGGCTCTGTGTTTGT	747		
Db	437	GACGTGTTTGGGCGAGGCGCACCGACGTGCTGCGCTCAAGAGGACCTGGGCTCTGTGTTTGT	496		
Qy	748	GACATCAAGTTCTCAGAGCTCTCCTATCTGTGAAGCCCTTCTGGGGCGACTACCTGAGTGGC	807		
Db	497	GACATCAAGTTCTCAGAGCTCTCCTATCTGTGAAGCCCTTCTGGGGCGACTACCTGAGTGGC	556		
Qy	808	GCCCTGTGAGAGGCGCCGCGGGGGGTGTTCTCTGACTAGAGCCCTGCGAGAGGCTGTGGGC	867		
Db	557	GCCCTGTGAGAGGCGCCCTGCGGGGGGTGTTCTCTGACTAGAGCCCTGCGAGAGGCTGTGGGC	616		
Qy	868	CGGAGAGCTGTGCGCTGCTGCTGCTCAGTGTGATGAGGCTGAC	909		
Db	617	CGGAGAGCTGTGCGCTGCTGCTGCTCAGTGTGATGAGGCTGAC	658		
RESULT 21					
BP306442		569 bp	mRNA	linear	EST 17-SEP-2004
LOCUS	BP306442				
DEFINITION	BP306442 Sugano cDNA library, macrophage Homo sapiens cDNA clone				
ACCESSION	BF007447				
VERSION	BP306442.1				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 569)				
JOURNAL	Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,				
	Mizushima-Sugano,J., Nakai,K. and Sugano,S.				
	Sequence comparison of human and mouse genes reveals a homologous				
	block structure in the promoter regions				
	Genome Res. 14 (9), 1711-1718 (2004)				

COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
SOURCE Location/Qualifiers
1. 569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="M807447"
/cell_type="macrophage"
/clone_lib="Sugano cDNA library, macrophage"

ORIGIN
Query Match 57.3%; Score 521; DB 5; Length 569;
Best Local Similarity 100.0%; Pred. No. 1.4e-253;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 GTGTCTCCAGAACGCTATAGCTATAGCACTTCAGACTTTCAGAGAGACAGAGGGTAC 387
DB 49 GTGTCTCCAGAACGCTATAGCTATAGCACTTCAGACTTTCAGAGAGACAGAGGGTAC 108
QY 388 TGCCGTGCGCGTCCGAGCTCAAGCAATTTCGAAATTTCAGAGGGTCAAGTGGAGACA 447
DB 109 TGCCGTGCGCGTCCGAGCTCAAGCAATTTCGAAATTTCAGAGGGTCAAGTGGAGACA 168
QY 448 GGCCTCCCCCAACCAAGCGGAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 507
DB 169 GGCCTCCCCCAACCAAGCGGAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 228
QY 508 CGGCGGCGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 567
DB 229 CGGCGGCGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 288
QY 568 TCTGAAGGCAAGTGAAGCTTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGCGAGCAT 627
DB 289 TCTGAAGGCAAGTGAAGCTTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGCGAGCAT 348
QY 628 GGGCGAGCGCTTGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 687
DB 349 GGGCGAGCGCTTGGAGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 408
QY 688 GACGTGTTTGGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 747
DB 409 GACGTGTTTGGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 468
QY 748 GACATCAAGTTCTGAGAGCTTCTCTATCTGAGCGCTTCTGAGGCGGCGGCGGCGG 807
DB 469 GACATCAAGTTCTGAGAGCTTCTCTATCTGAGCGCTTCTGAGGCGGCGGCGGCGG 528
QY 808 GCCCTGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 848
DB 529 GCCCTGCTGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 569

RESULT 22
BE883568 793 bp mRNA linear EST 20-OCT-2000
LOCUS 601508040P1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3909753 5',
DEFINITION mRNA sequence.
ACCESSION BE883568
VERSION BE883568.1 GI:10323244
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 793)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

EMAIL: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: LLM9723 row: P column: 10
High quality sequence stop: 678.
Location/Qualifiers
1. 793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3909753"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN
Query Match 56.7%; Score 515; DB 2; Length 793;
Best Local Similarity 100.0%; Pred. No. 1.5e-250;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTTTCAAAGAGACAGAGGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
DB 126 GCTTTCAAAGAGACAGAGGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 185
QY 422 ATTCTCAGCAGAGTCAAGTGGAGACAGGCTCCCCCAACCAAGGCGGAGCGGCGGAGTC 481
DB 186 ATTCTCAGCAGAGTCAAGTGGAGACAGGCTCCCCCAACCAAGGCGGAGCGGCGGAGTC 245
QY 482 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
DB 246 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 305
QY 542 AGCAGTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601
DB 306 AGCAGTCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 365
QY 602 GGGTTTCAGCAGAGTCAAGTGGAGACAGGCTCCCCCAACCAAGGCGGAGCTGCTGCTGCT 661
DB 366 GGGTTTCAGCAGAGTCAAGTGGAGACAGGCTCCCCCAACCAAGGCGGAGCTGCTGCTGCT 425
QY 662 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721
DB 426 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 485
QY 722 CAAGGACCTTGGGCTCTGTGTTTGTGACATCAAGTTCTCAGAGCTTCTCTATCTGAG 781
DB 486 CAAGGACCTTGGGCTCTGTGTTTGTGACATCAAGTTCTCAGAGCTTCTCTATCTGAG 545
QY 782 CCTTCTGGGCGGACCTACCTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841
DB 546 CCTTCTGGGCGGACCTACCTGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 605
QY 842 CTGAGGCGCTTGGAGAGGCTGTGGGCGGAGGCT 876
DB 606 CTGAGGCGCTTGGAGAGGCTGTGGGCGGAGGCT 640

RESULT 23
BE391120 654 bp mRNA linear EST 21-JUL-2000
LOCUS 601286078P1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3607843 5',
DEFINITION mRNA sequence.
ACCESSION BE391120
VERSION BE391120.1 GI:9336485
KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 654)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LICM259 row: 1 column: 20
 High quality sequence stop: 617.
 Location/Qualifiers
 1. 654

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3607843"
 /issue_type="endometrium, adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH MGC 44"
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-CDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 55.9%; Score 508; DB 2; Length 654;
 Best Local Similarity 100.0%; Pred. No. 5.6e-247;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 328 GTGCTCCAGAACGCTAGTATGAGCACTCCAGCTCTTCAAGAGAGACAGAGGTTAGC 387
Db 61 GTGCTCCAGAACGCTAGTATGAGCACTCCAGCTCTTCAAGAGAGACAGAGGTTAGC 120
Qy 388 TGCCGTGCGCGTGGCAGTCAAGCAATTCTGCAATTCAGCAGGTCAGTGGAGACA 447
Db 121 TGCCGTGCGCGTGGCAGTCAAGCAATTCTGCAATTCAGCAGGTCAGTGGAGACA 180
Qy 448 GGCTCCCCCAACCAAGCGGCGAGTGGGCGCGGCGGCGAGTGGTGCAGCA 507
Db 181 GGCTCCCCCAACCAAGCGGCGAGTGGGCGCGGCGGCGAGTGGTGCAGCA 240
Qy 508 CGGCGGGGAGAGGGGGCCCAAGCGGCGAGCCAGAGAGAGTGAAGCCCGGAGACTTCC 567
Db 241 CGGCGGGGAGAGGGGGCCCAAGCGGCGAGCCAGAGAGAGTGAAGCCCGGAGACTTCC 300
Qy 568 TCTGAAGGCAAAAGTACCTGTGATCATCCGGCTCCGGGTTGAGAGAGATCTGAGACAT 627
Db 301 TCTGAAGGCAAAAGTACCTGTGATCATCCGGCTCCGGGTTGAGAGAGATCTGAGACAT 360
Qy 628 GGGCGACCTTGGAGCAGAGGGGTGACATCCGGCGGGCCCAAGGCGCTGGCGGGCAGCTG 687
Db 361 GGGCGACCTTGGAGCAGAGGGGTGACATCCGGCGGGCCCAAGGCGCTGGCGGGCAGCTG 420
Qy 688 GACCTGTTGGGCGAGGCAACCGAGTGTGGCTCAAGGAGCCTGGGCTCTGTGTTTGT 747
Db 421 GACCTGTTGGGCGAGGCAACCGAGTGTGGCTCAAGGAGCCTGGGCTCTGTGTTTGT 480
Qy 748 GACATCAAGTTCTAGAGACTCTCTATCTGAGCGCTTCTGGGCGGACTACCTAGTGGC 807
Db 481 GACATCAAGTTCTAGAGACTCTCTATCTGAGCGCTTCTGGGCGGACTACCTAGTGGC 540

```

Qy 808 GCCCTGTGAGAGCCCTTGGCGGGCGTGT 835
 Db 541 GCCCTGTGAGAGCCCTTGGCGGGCGTGT 568

RESULT 24
 LOCUS BF655646 506 bp mRNA linear EST 21-DEC-2000
 DEFINITION 602119860F1 NIH_MGC_56 Homo sapiens CDNA clone IMAGE:4277031 5',
 mRNA sequence.
 ACCESSION BF655646
 VERSION BF655646.1 GI:11939541
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 506)
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LICM1098 row: 9 column: 16
 High quality sequence stop: 506.
 Location/Qualifiers
 1. 506

FEATURES
 source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4277031"
 /issue_type="primitive neuroectoderm"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NIH MGC 56"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggcgctcgagc); Site 2: SfiI (ggcattatggc);
 Double-stranded cDNA was prepared from cell line RNA. 5'
 and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCATTATGGC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCGCGGCGGCGGCGGCGGCGGAGT-
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 55.7%; Score 506; DB 2; Length 506;
 Best Local Similarity 100.0%; Pred. No. 5.9e-246;
 Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 301 CACCTGCGGCGAAGCGGCGCGGCGAGTGTCTCCAGAGCTATAGGACCTCC 360
Db 1 CACCTGCGGCGAAGCGGCGCGGCGAGTGTCTCCAGAGCTATAGGACCTCC 60
Qy 361 AGCTCTTCAAGAGAGAGAGGTTAGTGCCTGCGCGGCGAGTCAAGCAGTTCTGCA 420
Db 61 AGCTCTTCAAGAGAGAGAGGTTAGTGCCTGCGCGGCGAGTCAAGCAGTTCTGCA 120
Qy 421 AATTCTAGCAGAGGTCAGTGGAGACAGGCTCCCCCAACCAAGCGGCGGCGAGT 480
Db 121 AATTCTAGCAGAGGTCAGTGGAGACAGGCTCCCCCAACCAAGCGGCGGCGAGT 180
Qy 481 CGGGGCGGCGCGAGTGTGGTGCAGAGCGGCGGAGAGGGGGCCGAGCGGACCCGAG 540
Db 181 CGGGGCGGCGCGAGTGTGGTGCAGAGCGGCGGAGAGGGGGCCGAGCGGACCCGAG 240

```

QY 541 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAAAGTACCTGTGATCCGGCTC 600
DB 241 CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAAGCAAAAGTACCTGTGATCCGGCTC 300
QY 601 CGGGTTTGAGAGAGTACTGTCAGAGATGGGCCAGCTTTGAGAGAGGGGTGGCATCCCGG 660
DB 301 CGGGTTTGAGAGAGTACTGTCAGAGATGGGCCAGCTTTGAGAGAGGGGTGGCATCCCGG 360
QY 661 CGGGCCAGAGCGCTGGCGCGGCGAGCTGGACGTGTTTGGGCGAGGCCAGCGAGTGGCGC 720
DB 361 CGGGCCAGAGCGCTGGCGCGGCGAGCTGGACGTGTTTGGGCGAGGCCAGCGAGTGGCGC 420
QY 721 TCAAGGAGACCTGGGCTGTGTGTTGATCAATCAAGTTCTCAGAGCTTCTTATCTGGAC 780
DB 421 TCAAGGAGACCTGGGCTGTGTGTTGATCAATCAAGTTCTCAGAGCTTCTTATCTGGAC 480
QY 781 GCCCTTGGGGCGGCACTACCTGAGTGG 806
DB 481 GCCCTTGGGGCGGCACTACCTGAGTGG 506

RESULT 25

LOCUS CN373294 678 bp mRNA linear EST 16-MAY-2004
DEFINITION 1700059948013 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN373294
VERSION CN373294.1 GI:47373228
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 678)

AUTHORS

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Pisk, G.J.,
Li, Y., Xu, C., Feng, R., Guegler, K., Rao, M.S., Mandel, R.,
Lubkowsky, J. and Stanton, L.W.

TITLE

Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation

JOURNAL

Nat. Biotechnol. 22 (6), 707-716 (2004)

COMMENT

Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
220 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com

Insert Length: 678 Std Error: 0.00.
Location/Qualifiers

FEATURES

source
1..678
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, DMSO-treated H9 cell
line"
/clone_id="GRN_PREHEP"
/notes="Oligo dt primed, full-length enriched cDNA library
from DMSO-treated h9s cell line H9 (p22) maintained in
feeder-free conditions"

ORIGIN

Query Match 55.2%; Score 502; DB 7; Length 678;
Best Local Similarity 99.8%; Pred. No. 6.4e-244;
Matches 552; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 328 GTGTCTCCAGAACCTTATAGCTATGAGCAGCTCCAGCTTTTCAAGAGAGCAGAGGGTATGC 387
DB 88 GTGTCTCCAGAACCTTATAGCTATGAGCAGCTCCAGCTTTTCAAGAGAGCAGAGGGTATGC 147
QY 388 TGGCGTGGCGGTGGGAGGTCAAGAGTTCGCAAAATTCCTGAGCGGTCACTGGAGACA 447
DB 148 TGGCGTGGCGGTGGGAGGTCAAGAGTTCGCAAAATTCCTGAGCGGTCACTGGAGACA 207

QY 448 GGTCCCCCAACCAAGCGGAGCGGAGTGGGGCGGCGCCAGTGTGTGCCAGA 507
DB 208 GGTCCCCCAACCAAGCGGAGCGGAGTGGGGCGGCGCCAGTGTGTGCCAGA 267
QY 508 CGGGCGGAGAGGGGCGCCAGCGGACCCCAAGAGAGTATAGAGCCCGCCAGACTTTC 567
DB 268 CGGGCGGAGAGGGGCGCCAGCGGACCCCAAGAGAGTATAGAGCCCGCCAGACTTTC 327
QY 568 TCTGAAGCAAAAGTACCTGTGATCATCCGGCTCCGGGTTCTGAGAGAGTATCTGGAGCAT 627
DB 328 TCTGAAGCAAAAGTACCTGTGATCATCCGGCTCCGGGTTCTGAGAGAGTATCTGGAGCAT 387
QY 628 GGGCGAGCTTTGAGCAAGGCGGTGGCATCCGGCGGCGCCAGCGGTGGCGCGAGCTGG 687
DB 388 GGGCGAGCTTTGAGCAAGGCGGTGGCATCCGGCGGCGCCAGCGGTGGCGCGAGCTGG 447
QY 688 GACGTGTTTGGGAGGCGCACCGCAGTGTGGGCTCAAGAGGACCTGGGCTGTGTGTGTGT 747
DB 448 GACGTGTTTGGGAGGCGCACCGCAGTGTGGGCTCAAGAGGACCTGGGCTGTGTGTGTGT 507
QY 748 GACATCAAGTTCTCAGAGCTTCTCTATCTGAGACGCTTTCTGGGCGAGTACTTGAAGTGGC 807
DB 508 GACATCAAGTTCTCAGAGCTTCTCTATCTGAGACGCTTTCTGGGCGAGTACTTGAAGTGGC 567
QY 808 GCCCTGTGACAGGCGCTGCGGGGCGGTCTTCTGATGAGAGCGCTTGGAGAGGCTGTGGGC 867
DB 568 GCCCTGTGACAGGCGCTGCGGGGCGGTCTTCTGATGAGAGCGCTTGGAGAGGCTGTGGGC 627
QY 868 CGGAGGCTGTTC 880
DB 628 CGGAGGCTGTTC 640

RESULT 26

LOCUS B0923358 931 bp mRNA linear EST 20-AUG-2002
DEFINITION AGENCOURT 8784534 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6376221
5', mRNA sequence.
ACCESSION B0923358
VERSION B0923358.1 GI:22338389
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 931)

AUTHORS

NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

COMMENT

Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LNCM2558 row: a column: 22
High quality sequence stop: 511.
Location/Qualifiers

FEATURES

source
1..931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6376221"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_43"
/notes="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into Scori/XhoI sites using the following 5'
adaptor: GGCAAGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. | "

ORIGIN

Query Match 55.0%; Score 500; DB 5; Length 931;
Best Local Similarity 99.8%; Pred. No. 6.6e-243;
Matches 550; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGGCGCTATCCGGGTGAGCCCGCGCCGCTGGCGGAGAGATGAGTCCCTGGACTAC 60
DB 10 ATGGCGCTATCCGGGTGAGCCCGCGCCGCTGGCGGAGAGATGAGTCCCTGGACTAC 69
QY 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGCACTGACCGAGTGC 120
DB 70 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGCACTGACCGAGTGC 129
QY 121 GAGCTGAGTCTCTGGCTTTCTGCTGATGAGGCTCTCGCGCCGCGAGGCTTAC 180
DB 130 GAGCTGAGTCTCTGGCTTTCTGCTGATGAGGCTCTCGCGCCGCGAGGCTTAC 189
QY 181 CGGGCCCGGAGCGGCTGAGGCTCTGCTGAGGCTGAGGCGCGCGGCACTGAGG 240
DB 190 CGGGCCCGGAGCGGCTGAGGCTCTGCTGAGGCTGAGGCGCGCGGCACTGAGG 249
QY 241 AGCAACCTGCGGTGCTGCGGCACTCTCGCGCTGCTGCGCCGCGCACTGCTCCG 300
DB 250 AGCAACCTGCGGTGCTGCGGCACTCTCGCGCTGCTGCGCCGCGCACTGCTCCG 309
QY 301 CACCTGGCGGCAAGCGGCGCGGCACTGCTTCAAGAGCTTACTATGGCACTCC 360
DB 310 CACCTGGCGGCAAGCGGCGCGGCACTGCTTCAAGAGCTTACTATGGCACTCC 369
QY 361 AGCTCTCAAGAGAGAGAGAGGAGTGTGCTGCGGCTGCGGCACTGAGGCTTCA 420
DB 370 AGCTCTCAAGAGAGAGAGAGGAGTGTGCTGCGGCTGCGGCACTGAGGCTTCA 429
QY 421 AATTCTCAGAGAGGTGAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
DB 430 AATTCTCAGAGAGGTGAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 489
QY 481 CGGGGCGGCGGCACTGTGTGCTGCGAGCGGCGGAGAGAGGCGCCCAAGCGG 540
DB 490 CGGGGCGGCGGCACTGTGTGCTGCGAGCGGCGGAGAGAGGCGCCCAAGCGG 549
QY 541 CAGAGTCAGA 551
DB 550 CAGAGTCAGA 560

```

RESULT 27
BU602002 783 bp mRNA linear EST 20-SEP-2002
LOCUS
DEFINITION ABENCOURT_10016659 NIH_MGC_142 Homo sapiens cDNA clone
IMAGE:6496027 5', mRNA sequence.
ACCESSION BU602002
VERSION BU602002.1 GI:23253761
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: NCI
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2674 row: m column: 20
High quality sequence stop: 558.
Location/Qualifiers

FEATURES

Source

```

1..783
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6496027"
/issue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (TI-phage-resistant)"
/clone_lib="NIH_MGC_142"
/notes="vector: pDNR-LIB; Site 1: SfiI (ggccatcagcc); Site 2: SfiI (ggcgctcgccg); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTCAATCAACGACAGTGGCCATTAACGGCGG-3' and
5'-ATTCTAGAGCCGAGGCGCCGACATG-dt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC_141). Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match 52.8%; Score 480; DB 5; Length 783;
Best Local Similarity 99.7%; Pred. No. 1e-232;
Matches 580; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 328 GTGTCTCAAGAGCTTACTATGAGTCCAGCTCTTCAAGAGAGAGAGGCTAGC 387
DB 64 GTGTCTCAAGAGCTTACTATGAGTCCAGCTCTTCAAGAGAGAGAGGCTAGC 123
QY 388 TGCCTGCGCGGTGCGGAGTCAAGAGTTCGAAATTCAGCAGGCTCAGTGGAGACA 447
DB 124 TGCCTGCGCGGTGCGGAGTCAAGAGTTCGAAATTCAGCAGGCTCAGTGGAGACA 183
QY 448 GGTCTCCCCCAACCAAGCGGAGCGGAGTGTGGGCGCGCCCAAGTGTGTCCAGA 507
DB 184 GGTCTCCCCCAACCAAGCGGAGCGGAGTGTGGGCGCGCCCAAGTGTGTCCAGA 243
QY 508 CGGCGGCGAGAGAGGCGCCCAAGCGGAGCGGAGTGTGGGCGCGCCCAAGTGT 567
DB 244 CGGCGGCGAGAGAGGCGCCCAAGCGGAGCGGAGTGTGGGCGCGCCCAAGTGT 303
QY 568 TCTGAAGCAAGTGAACCTGTGACATCGGCTCGGGTGTGAGAGAGTACTGAGACAT 627
DB 304 TCTGAAGCAAGTGAACCTGTGACATCGGCTCGGGTGTGAGAGAGTACTGAGACAT 363
QY 628 GGGCGACCTTGGAGCAAGGCGGTGATCCCGGCGGCCCAAGGCGCTGGCGGCACTG 687
DB 364 GGGCGACCTTGGAGCAAGGCGGTGATCCCGGCGGCCCAAGGCGCTGGCGGCACTG 423
QY 688 GAGCTTTTGGGAGGAGGACCGGAGTGTGCTGCTCAAGAGGAGCTGGGCTCTGTGTGT 747
DB 424 GAGCTTTTGGGAGGAGGACCGGAGTGTGCTGCTCAAGAGGAGCTGGGCTCTGTGTGT 483
QY 748 GACATCAAGTTCAGAGCTCTCTATCTGAGAGCTTGTGGGCGACTACTAGTGGC 807
DB 484 GACATCAAGTTCAGAGCTCTCTATCTGAGAGCTTGTGGGCGACTACTAGTGGC 543
QY 808 GCCCTGTGAGGCGCTGCGGCGGTCTCTGACTGAGGCGCTGCGAGAGGCTGTGGGC 867
DB 544 GCCCTGTGAGGCGCTGCGGCGGTCTCTGACTGAGGCGCTGCGAGAGGCTGTGGGC 603
QY 868 CGGAGGCTGTTGCTGCTGTGATGTGATGAGGCTGAC 909

```



```
QY 124 CTGAGCTCTGGCTTTCTGCTGGATGAGGCTCTGGCCGCGGAGGCTTACCCCG 183
Db 210 CTGAGCTCTGGCTTTCTGCTGGATGAGGCTCTGGCCGCGGAGGCTTACCCCG 269
QY 184 GCCCGCAGCGGCTAGAGCTCTGCTGAGAGCGCGCGGAGAGTTCGAGAGC 243
Db 270 GCCCGCAGCGGCTTAGAGCTCTGCTGAGAGCGCGCGGAGAGTTCGAGAGC 329
QY 244 AACCTGCGCTGCTGGGCGCACTCTGCGCGTGTGGCCCGCGCAGACTGTCCGCG 303
Db 330 AACCTGCGCTGCTGGGCGCACTCTGCGCGTGTGGCCCGCGCAGACTGTCCGCG 389
QY 304 CTGGCGCGCAGCGCGCGCGCGCGAGTGTCTTCAGAAAGCTATAGCACTTCCAGC 363
Db 390 CTGGCGCGCAGCGCGCGCGCGCGAGTGTCTTCAGAAAGCTATAGCACTTCCAGC 449
QY 364 TCTTCAAGAGAGCAGAGGGTGTGCTGCGCTGCGCGAGTCAAGAGATTGTGCAAT 423
Db 450 TCTTCAAGAGAGCAGAGGGTGTGCTGCGCTGCGCGAGTCAAGAGATTGTGCAAT 509
QY 424 TCTCAGCAGAGTCAAGTGGAGACAGGCTCCCCCAACCAAGCGCAGCGCGAGTCG 483
Db 510 TCTCAGCAGAGTCAAGTGGAGACAGGCTCCCCCAACCAAGCGCAGCGCGAGTCG 569
QY 484 GCGCGCGCCAGTGTGTGTCAGACGCGCGCGAGAGGGGCCCAAGCCGACCCAGCAG 543
Db 570 GCGCGCGCCAGTGTGTGTCAGACGCGCGCGAGAGGGGCCCAAGCCGACCCAGCAG 629
QY 544 CAGTCAGAGCCCGCAGACCTTCTCTGAAGCAAGTGACC 585
Db 630 CAGTCAGAGCCCGCAGACCTTCTCTGAAGCAAGTGACC 671

RESULT 30
CN373291 577 bp mRNA linear EST 16-MAY-2004
LOCUS 17000425073610 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN373291
VERSION CN373291.1 GI:47373225
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 577)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalam, R.,
Lebkowski, J. and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 577 Std Error: 0.00.
FEATURES
Source
1..577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="Oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."
```

```
Query Match 50.4%; Score 458; DB 7; Length 577;
Best Local Similarity 99.8%; Pred. No. 1,6e-221;
Matches 508; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 372 GAGGACAGAGGAGTAGTCCGTCGCCGTGGAGTCAAGAGTTCGCAATTCTCAGCA 431
Db 1 GAGGACAGAGGAGTAGTCCGTCGCCGTGGAGTCAAGAGTTCGCAATTCTCAGCA 60
QY 432 GGGTCACTGGAGACAGAGCTTCCCCCAACCAAGCGCAGCGCGAGTGGGGCCGCC 491
Db 61 GGGTCACTGGAGACAGAGCTTCCCCCAACCAAGCGCAGCGCGAGTGGGGCCGCC 120
QY 492 CAGTGTGTGTCCAGAGCGCGCGAGAGAGGGGCCCGCGCGCAGCGAGTGAAGA 551
Db 121 CAGTGTGTGTCCAGAGCGCGCGAGAGAGGGGCCCGCGCGCAGCGAGTGAAGA 180
QY 552 GCCCGCAGACCTTCTCTGAAGCAAGTGAATCTGATCATCGGCTCCGGGTTGAGC 611
Db 181 GCCCGCAGACCTTCTCTGAAGCAAGTGAATCTGATCATCGGCTCCGGGTTGAGC 240
QY 612 AGAGTACTGGAGATGGGCAAGCTTTGAGAGCAGGGGCTGCAATCCGGCGGCCAGGC 671
Db 241 AGAGTACTGGAGATGGGCAAGCTTTGAGAGCAGGGGCTGCAATCCGGCGGCCAGGC 300
QY 672 GCTGGCGCGGACAGTGAAGCTGTTGGGCAAGCCACCGCAGTGTGGCTCAAGGACCT 731
Db 301 GCTGGCGCGGACAGTGAAGCTGTTGGGCAAGCCACCGCAGTGTGGCTCAAGGACCT 360
QY 732 GGGCTGTGTGTTGTGATCAATCAAGTCTCAGAGCTCTCTATCTGAGACCTTCTGGGG 791
Db 361 GGGCTGTGTGTTGTGATCAATCAAGTCTCAGAGCTCTCTATCTGAGACCTTCTGGGG 420
QY 792 CGACTACCTGATGAGCGCCCTGTGACAGGCCCTGGCGGCGTGTCTGACTGAGGCCCT 851
Db 421 CGACTACCTGATGAGCGCCCTGTGACAGGCCCTGGCGGCGTGTCTGACTGAGGCCCT 480
QY 852 GCGAGAGCTGTGGCGCGCGAGAGGCTGTC 880
Db 481 GCGAGAGCTGTGGCGCGCGAGAGGCTGTC 509

RESULT 31
BQ956773 991 bp mRNA linear EST 21-AUG-2002
LOCUS BQ956773
DEFINITION AGENCOURT 8804443 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6377779
5', mRNA sequence.
ACCESSION BQ956773
VERSION BQ956773.1 GI:22372251
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 991)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@ncl.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LWM2562 row: b column: 20
High quality sequence stop: 542.
FEATURES
Source
1..991
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```


/clone="IMAGE:6377779"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_47"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.9%; Score 454; DB 5; Length 991;
Best Local Similarity 100.0%; Pred. No. 1.7e-219; Mismatches 454; Conservative 0; Indels 0; Gaps 0;
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 456 CCCAACCAAGCGGAGCGGCGGAGTCCGGGCGCGCCAGTGTGTCACAGCGGCGG 515
Db 1 CCCAACCAAGCGGAGCGGCGGAGTCCGGGCGCGCCAGTGTGTCACAGCGGCGG 60
Qy 516 GAGAGGGGCGCGGCGGAGCGGCGGAGTCCGGGCGCGCCAGTGTGTCACAGCGG 575
Db 61 GAGAGGGGCGCGGCGGAGCGGCGGAGTCCGGGCGCGCCAGTGTGTCACAGCGG 120
Qy 576 CAAAGTACCTGTACATCCGGGCTCCGGGTTCCAGACAGTACAGTACAGTACAGT 635
Db 121 CAAAGTACCTGTACATCCGGGCTCCGGGTTCCAGACAGTACAGTACAGTACAGT 180
Qy 636 CTTGAGACAGGCGGTGCGATCCGGGCGCGCCAGGCGGTGCGCGGAGTGAAGTGT 695
Db 181 CTTGAGACAGGCGGTGCGATCCGGGCGCGCCAGGCGGTGCGCGGAGTGAAGTGT 240
Qy 696 TGGGAGGCGGCGGCGGAGTGTGCGCTCAAGGAGCTTGGGCTCTGTGTTTGACATCA 755
Db 241 TGGGAGGCGGCGGCGGAGTGTGCGCTCAAGGAGCTTGGGCTCTGTGTTTGACATCA 300
Qy 756 GTTCTCAGAGCTCTTCATCTGAGAGCGCTTCTGGGAGCACTACCTGAGTGGCGCTT 815
Db 301 GTTCTCAGAGCTCTTCATCTGAGAGCGCTTCTGGGAGCACTACCTGAGTGGCGCTT 360
Qy 816 GCAGAGCGCTGCGGCGGCGGTCTGACTGAGAGCGCTTGGAGGCTGTGGGCGGAGAGC 875
Db 361 GCAGAGCGCTGCGGCGGCGGTCTGACTGAGAGCGCTTGGAGGCTGTGGGCGGAGAGC 420
Qy 876 TGTTCGCTGCTGTGCTGATGTGATGAGGCTGAC 909
Db 421 TGTTCGCTGCTGTGCTGATGTGATGAGGCTGAC 454
RESULT 32
LOCUS B1827768 832 bp mRNA linear EST 04-OCT-2001
DEFINITION 60307430F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167254 5',
mRNA sequence.
ACCESSION B1827768 GI:15939318
VERSION B1827768.1
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 832)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bld-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: L14M1415 row: 1 column: 07
High quality sequence stop: 670.
Location/Qualifiers

FEATURES
source

1..832
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5167254"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dt primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.8%; Score 453; DB 4; Length 832;
Best Local Similarity 99.8%; Pred. No. 5.6e-219; Mismatches 503; Conservative 0; Indels 1; Gaps 0;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGCGGCTATCCGGGCTGACCCCGGCGCGGCTGTGGAGAGATGAGTGTGCTGACTAC 60
Db 65 ATGCGGCTATCCGGGCTGACCCCGGCGCGGCTGTGGAGAGATGAGTGTGCTGACTAC 124
Qy 61 TACGGGATGCTGTGCTTACCGTATGTTGAGGTGTGGCGGCACTGACGAGTGC 120
Db 125 TACGGGATGCTGTGCTTACCGTATGTTGAGGTGTGGCGGCACTGACGAGTGC 184
Qy 121 GAGCTGAGAGCTCTGGGCTTCTGCTGGATGAGGCTCTGGGCGCGGAGGCTTAGCC 180
Db 185 GAGCTGAGAGCTCTGGGCTTCTGCTGGATGAGGCTCTGGGCGCGGAGGCTTAGCC 244
Qy 181 CGGGCCGCGAGCGGCTAGAGAGCTCTGTGAGCTGAGAGCGCGGCGGCACTGCGGAG 240
Db 245 CGGGCCGCGAGCGGCTAGAGAGCTCTGTGAGCTGAGAGCGCGGCGGCACTGCGGAG 304
Qy 241 AGCAACTCGCGGCTGTGGGCAACTCTGCGGCTGTGCGCCGCACTGCTGCGG 300
Db 305 AGCAACTCGCGGCTGTGGGCAACTCTGCGGCTGTGCGCCGCACTGCTGCGG 364
Qy 301 CACCTGGGCGGCGGCGGCGGCGGCACTGCTCAGAAAGCTATGAGTATGAGCACTCC 360
Db 365 CACCTGGGCGGCGGCGGCGGCGGCACTGCTCAGAAAGCTATGAGTATGAGCACTCC 424
Qy 361 AGCTCTTCAAGAGAGAGAGAGGTAGTCCGCTGCGGCTGCGGAGTCAAGAGTTCTGCA 420
Db 425 AGCTCTTCAAGAGAGAGAGGTAGTCCGCTGCGGCTGCGGAGTCAAGAGTTCTGCA 484
Qy 421 AATTCTCAGAGGCTCAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGGAGT 480
Db 485 AATTCTCAGAGGCTCAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGGAGT 544
Qy 481 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 504
Db 545 CGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 568
RESULT 33
LOCUS BM928759 1028 bp mRNA linear EST 12-MAR-2002
DEFINITION BM928759 6726912 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5798907
5', mRNA sequence.
ACCESSION BM928759

VERSION BM928759.1 GI:19379138
 EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1028)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 CDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM2027 row: k column: 04
 High quality sequence stop: 585.
 Location/Qualifiers
 1. 1028
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5798907"
 /issue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_100"
 /note="Organ: Liver; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 49.5%; Score 450; DB 5; Length 1028;
 Best Local Similarity 100.0%; Pred. No. 1.9e-217;
 Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 460 ACCAAGCGGAGCGGCGGAGTCGGGCGGCCAGTGTGTGTCAGACGGCGCGAGA 519
 1 ACCAAGCGGAGCGGCGGAGTCGGGCGGCCAGTGTGTGTCAGACGGCGCGAGA 60
 520 GGGGCCCCAGCGGAGCGGCGGAGTCAGAGCCGCGCAGACCTTCTTGAAGCAAA 579
 61 GGGGCCCCAGCGGAGCGGCGGAGTCAGAGCCGCGCAGACCTTCTTGAAGCAAA 120
 580 GTGACCTGTACATCCGGCTCCGGGTTGAGCAGAGTACGCGAGCANTGGCCGAGCTTG 639
 121 GTGACCTGTACATCCGGCTCCGGGTTGAGCAGAGTACGCGAGCANTGGCCGAGCTTG 180
 640 GAGCAGGCGGTGATCCGGCGGCGCCAGAGCGCTGGCGCGGAGCTGAGTGTGGG 699
 181 GAGCAGGCGGTGATCCGGCGGCGCCAGAGCGCTGGCGCGGAGCTGAGTGTGGG 240
 700 CAGGCAACCGAGTGTGCGCTCAAGGACCTGGGCTCTGTGTGTGTGATCAAGTTTC 759
 241 CAGGCAACCGAGTGTGCGCTCAAGGACCTGGGCTCTGTGTGTGTGATCAAGTTTC 300
 760 TCAGAGCTTCTCTATCTGAGAGCTTCTGTGGGCGACATACCTGATGGCGCCCTGCTCAG 819
 301 TCAGAGCTTCTCTATCTGAGAGCTTCTGTGGGCGACATACCTGATGGCGCCCTGCTCAG 360
 820 GCCCTGCGGGCGGTCTCTGACTGAGAGCCCTGCGAGAGGCTGTGGCCGGGAGGCTGTT 879
 361 GCCCTGCGGGCGGTCTCTGACTGAGAGCCCTGCGAGAGGCTGTGGCCGGGAGGCTGTT 420

QY 880 CGCCTGCTGCTACGTGTGATGAGGCTGAC 909
 Db 421 CGCCTGCTGCTACGTGTGATGAGGCTGAC 450
 RESULT 34
 LOCUS BM927715
 DEFINITION AGENCOURT_6728785 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:5797541
 5', mRNA sequence.
 ACCESSION BM927715
 VERSION BM927715.1 GI:19378094
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1033)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LNCM2024 row: b column: 06
 High quality sequence stop: 498.
 Location/Qualifiers
 1. 1033
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5797541"
 /issue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_100"
 /note="Organ: Liver; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

FEATURES

source
 1. 1033
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5797541"
 /issue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_100"
 /note="Organ: Liver; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

ORIGIN

Query Match 49.5%; Score 450; DB 5; Length 1033;
 Best Local Similarity 100.0%; Pred. No. 1.9e-217;
 Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 460 ACCAAGCGGAGCGGCGGAGTCGGGCGGCCAGTGTGTGTCAGACGGCGCGAGA 519
 1 ACCAAGCGGAGCGGCGGAGTCGGGCGGCCAGTGTGTGTCAGACGGCGCGAGA 60
 520 GGGGCCCCAGCGGAGCGGCGGAGTCAGAGCCGCGCAGACCTTCTTGAAGCAAA 579
 61 GGGGCCCCAGCGGAGCGGCGGAGTCAGAGCCGCGCAGACCTTCTTGAAGCAAA 120
 580 GTGACCTGTACATCCGGCTCCGGGTTGAGCAGAGTACGCGAGCANTGGCCGAGCTTG 639
 121 GTGACCTGTACATCCGGCTCCGGGTTGAGCAGAGTACGCGAGCANTGGCCGAGCTTG 180
 640 GAGCAGGCGGTGATCCGGCGGCGCCAGAGCGCTGGCGCGGAGCTGAGTGTGGG 699
 181 GAGCAGGCGGTGATCCGGCGGCGCCAGAGCGCTGGCGCGGAGCTGAGTGTGGG 240
 700 CAGGCAACCGAGTGTGCGCTCAAGGACCTGGGCTCTGTGTGTGTGATCAAGTTTC 759

```
Db      241 CAGGCCACCCGACGCTGCGCCCTCAAGGAGACCTGGGCTCTGTGTGATCATCAAGTTC 300
Qy      760 TCAGAGCTCTCTATCTGAGAGCCTTCTGGGGGAGCACTACCTGATGGGCGCCCTGCGAG 819
Db      301 TCAGAGCTCTCTATCTGAGAGCCTTCTGGGGGAGCACTACCTGATGGGCGCCCTGCGAG 360
Qy      820 GCCCTGGGGGCGCTTCTGACTGAGGCGCTGGGAGAGGCTGTGGCCGGGAGGCTGTT 879
Db      361 GCCCTGGGGGCGGTTCTGACTGAGGCGCTGGGAGAGGCTGTGGCCGGGAGGCTGTT 420
Qy      880 CGCCTGCTGTGATGTGATGAGGCTGAC 909
Db      421 CGCCTGCTGTGATGTGATGAGGCTGAC 450

RESULT 35
BP343733 586 bp mRNA linear EST 17-SEP-2004
LOCUS    BP343733 Sugano cDNA library, stomach mucosa Homo sapiens cDNA
DEFINITION
ACCESSION BP343733
VERSION    BP343733.1 GI:52273660
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS   Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
JOURNAL    block structure in the promoter regions
COMMENT     Genome Res. 14 (9), 1711-1718 (2004)
CONTACT:   Yutaka Suzuki
DEPARTMENT of Virology
INSTITUTE of Medical Science, University of Tokyo
4-6-1, Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES   Location/Qualifiers
source     1..586
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="STM08261"
            /tissue_type="stomach mucosa"
            /clone_id="Sugano cDNA library, stomach mucosa"

ORIGIN
Query Match 49.1%; Score 446; DB 5; Length 586;
Best Local Similarity 99.8%; Pred. No. 2,1e-215;
Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
Db      85 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 144
Qy      61 TAAGGATGCTGTGCTTCAACCGTATGTGAGAGTGTGGCGGGCAACTGACCGAGTGC 120
Db      145 TAAGGATGCTGTGCTTCAACCGTATGTGAGAGTGTGGCGGGCAACTGACCGAGTGC 204
Qy      121 GAGCTGAGAGTCTGAGCTTCTGCTGATGAGGCTCTGGCGCGCGCGAGGCTTATGCC 180
Db      205 GAGCTGAGAGTCTGAGCTTCTGCTGATGAGGCTCTGGCGCGCGCGAGGCTTATGCC 264
Qy      181 CGGGCCCGCAGCGGCTTGAAGCTCTGTGAGAGTGAAGCGCGCGGAGTGGCGGAG 240
Db      265 CGGGCCCGCAGCGGCTTGAAGCTCTGTGAGAGTGAAGCGCGCGGAGTGGCGGAG 324
Qy      241 AGCAACTGTGGGCTGTGGGCAATCTCTGGCGCTGTGGCGCGGCAAGACTGTGCGG 300
Db      325 AGCAACTGTGGGCTGTGGGCAATCTCTGGCGCTGTGGCGCGGCAAGACTGTGCGG 384
```

```
Qy      301 CACCTGCGCGCAGCGCGCCCGCCAGTGTCTTCAGAGGCTATGATGACCTCC 360
Db      385 CACCTGCGCGCAGCGCGCCCGCCAGTGTCTTCAGAGGCTATGATGACCTCC 444
Qy      361 AGCTCTCAAGAGAGAGAGGATAGTGCCTGTGGCGGAGTCAAGAGTTCGCA 420
Db      445 AGCTCTCAAGAGAGAGAGGATAGTGCCTGTGGCGGAGTCAAGAGTTCGCA 504
Qy      421 AATTCTCAGAGAGGTCACTGGGAGACAGAGCTCCCGCCCAACCAAGCGGCGGAGT 480
Db      505 AATTCTCAGAGAGGTCACTGGGAGACAGAGCTCCCGCCCAACCAAGCGGCGGAGT 564
Qy      481 CGGGCCCGCCCACTGG 497
Db      565 CGGGCCCGCCCACTGG 581

RESULT 36
CB994488 780 bp mRNA linear EST 01-MAY-2003
LOCUS    CB994488
DEFINITION
ACCESSION CB994488
VERSION    CB994488.1 GI:30289008
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 780)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strauberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
            cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
            cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: NDAM351 row: p column: 13
            High quality sequence stop: 576.
            Location/Qualifiers
            1..780
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:30333396"
            /tissue_type="pre-ecialupltic placenta"
            /lab_host="DH10B TONA"
            /clone_id="NIH MGC 148"
            /notes="Organ: placenta; Vector: pBluescriptR; Site 1:
            all-XhoI; Site 2: BamH; Library: is oligo-dT primed and
            directionally cloned using primer
            5'-TTTCTTTTCTTTTCTTTN-3', size-selected for average insert
            size 2.3 kb and normalized to R07 5. This is a primary
            library enriched for full-length clones and constructed
            using the Cap-trapper method (Carninci, in preparation).
            Library constructed by M. Brownstein (NIH/NHGRI,
            National Institutes of Health). Note: this is a NIH-MGC
            Library."
```

```
ORIGIN
Query Match 49.1%; Score 446; DB 6; Length 780;
Best Local Similarity 99.6%; Pred. No. 2,1e-215;
Matches 546; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
Db      117 ATGGCGCTATCCGGGTGACCCCGCCCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 176
```

```
QY 61 TACGGATGCTGTCGCTTACCCGTATGTTGAGGTGTGGGCGGGCAACTGACCGAGTGC 120
    |||
Db 177 TACGGATGCTGTCGCTTACCCGTATGTTGAGGTGTGGGCGGGCAACTGACCGAGTGC 236
QY 121 GACCTGAGAGCTCTGGGCTTTCTGCTGGATGAGAGGCTCTGGGCGCGCGGAGGCTTAAGCC 180
    |||
Db 237 GACCTGAGAGCTCTGGGCTTTCTGCTGGATGAGAGGCTCTGGGCGCGCGGAGGCTTAAGCC 236
QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCTGAGCGCGCGGAGGCTAGTGCAGG 240
    |||
Db 297 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCTGAGCGCGCGGAGGCTAGTGCAGG 356
QY 241 AGCAACCTGGGCTGCTGGGGCAACTCTGCGGCTGTGGCCCGGCAAGCTGTGCGG 300
    |||
Db 357 AGCAACCTGGGCTGCTGGGGCAACTCTCTGCGGCTGTGGCCCGGCAAGCTGTGCGG 416
QY 301 CACCTGGCGGCAAGCGGCGCGCGGCAAGTCTTCAGAAAGCTATAGCTATGCACTCC 360
    |||
Db 417 CACCTGGCGGCAAGCGGCGCGCGGCAAGTCTTCAGAAAGCTATAGCTATGCACTCC 476
QY 361 AGCTCTTCAAGAAGAGCAGAGGGTAGCTGCGCTGCGCGGAGTCAAGCAGTTCTGCA 420
    |||
Db 477 AGCTCTTCAAGAAGAGCAGAGGGTAGCTGCGCTGCGCGGAGTCAAGCAGTTCTGCA 536
QY 421 AATTCTCAGAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
    |||
Db 537 AATTCTCAGAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 536
QY 481 CGGGGCGGCGGCAAGTGTGTGTCAGAGCGGCGGAGGAGGAGGCGGAGCGGAG 540
    |||
Db 597 CGGGGCGGCGGCAAGTGTGTGTCAGAGCGGCGGAGGAGGAGGCGGAGCGGAG 656
QY 541 CACGACATC 548
    |||
Db 657 CACGACATC 664

RESULT 37
CB992627 831 bp mRNA linear EST 01-MAY-2003
LOCUS AGENCOURT 13645069 NIH_MGC_148 Homo sapiens cDNA clone
DEFINITION IMAGE:30339575 5', mRNA sequence.
ACCESSION CB992627
VERSION CB992627.1 GI:30287147
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabds-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: NDAM368 row: a column: 24
High quality sequence stop: 525.
Location/Qualifiers
1. 831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30339575"
/tissue_type="pre-eclamptic placenta"
```

ORIGIN

```
Query Match 48.6%; Score 442; DB 6; Length 831;
Best Local Similarity 99.8%; Pred. No. 2.2e-213;
Matches 492; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTGCAACCGCGCCGCTGCTGGAGAGATGATGCTTGAATAC 60
    |||
Db 117 ATGGCGCTATCCGGGTGCAACCGCGCCGCTGCTGGAGAGATGATGCTTGAATAC 176
QY 61 TACGGATGCTGTCGCTTACCCGTATGTTGAGGTGTGGGCGGGCAACTGACCGAGTGC 120
    |||
Db 177 TACGGATGCTGTCGCTTACCCGTATGTTGAGGTGTGGGCGGGCAACTGACCGAGTGC 236
QY 121 GACCTGAGAGCTCTGGGCTTTCTGCTGGATGAGAGCTCTGGGCGCGGAGGCTTAAGCC 180
    |||
Db 237 GACCTGAGAGCTCTGGGCTTTCTGCTGGATGAGAGCTCTGGGCGCGGAGGCTTAAGCC 296
QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCTGAGCGCGCGGAGGCTAGTGCAGG 240
    |||
Db 297 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGCTGAGCGCGCGGAGGCTAGTGCAGG 356
QY 241 AGCAACCTGGGCTGCTGGGGCAACTCTCTGCGGCTGTGGCCCGGCAAGCTGTGCGG 300
    |||
Db 357 AGCAACCTGGGCTGCTGGGGCAACTCTCTGCGGCTGTGGCCCGGCAAGCTGTGCGG 416
QY 301 CACCTGGCGGCAAGCGGCGCGCGGCAAGTCTTCAGAAAGCTATAGCTATGCACTCC 360
    |||
Db 417 CACCTGGCGGCAAGCGGCGCGCGGCAAGTCTTCAGAAAGCTATAGCTATGCACTCC 476
QY 421 AATTCTCAGAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
    |||
Db 537 AATTCTCAGAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 536
QY 481 CGGGGCGGCGGCAAGTGTGTGTCAGAGCGGCGGAGGAGGAGGCGGAGCGGAG 540
    |||
Db 597 CGGGGCGGCGGCAAGTGTGTGTCAGAGCGGCGGAGGAGGAGGCGGAGCGGAG 656

RESULT 38
B0072022 1042 bp mRNA linear EST 02-APR-2002
LOCUS AGENCOURT 6859739 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5928508
DEFINITION 5', mRNA sequence.
ACCESSION B0072022
VERSION B0072022.1 GI:19901068
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabds-remail.nih.gov
```

Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.gov>
Plate: L1CM2105 row: c column: 05
High quality sequence stop: 620.
Location/Qualifiers

FEATURES
SOURCE

```
1. 1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5928508"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_47"
/note="Organ: brain; Vector: pOT7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G) Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match 48.4%; Score 440; DB 5; Length 1042;
Best Local Similarity 99.8%; Pred. No. 2.3e-212;
Matches 490; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 114 CGAGTGGAGGCTGAGGCTCTGCGCTTCTGCTGATGAGGCTCTGCGCGCCGCGGAG 173
    |||
Db 212 CGAGTGGAGGCTGAGGCTCTGCGCTTCTGCTGATGAGGCTCTGCGCGCCGCGGAG 271

QY 174 CTTAGCCCGGCGCGGAGGCTGAGGCTCTGCGCTTCTGCGCTGAGGCTGAGGCGCGGAG 233
    |||
Db 272 CTTAGCCCGGCGCGGAGGCTGAGGCTCTGCGCTTCTGCGCTGAGGCTGAGGCGCGGAG 331

QY 234 CGGCGAGAGCAACCTGCGGCTGCTGCGGCAACTCTGCGCTGCTGCGCGCGGAGCCT 293
    |||
Db 332 CGAGGAGAGCAACCTGCGGCTGCTGCGGCAACTCTGCGCTGCTGCGCGCGGAGCCT 391

QY 294 GCTCCCGGAGCTGCGGCGGAGGCTGCGGCGGCGGCGGAGGCTGCTGCGGCGGAGCAG 353
    |||
Db 392 GCTCCCGGAGCTGCGGCGGAGGCTGCGGCGGCGGCGGAGGCTGCTGCGGCGGAGCAG 451

QY 354 CACCTCCAGCTCTTCAAGAAGAGAGGAGTGTGCGGCTGCGGCGGAGTCAAGCAG 413
    |||
Db 452 CACCTCCAGCTCTTCAAGAAGAGAGGAGTGTGCGGCTGCGGCGGAGTCAAGCAG 511

QY 414 TTCTGCAAAATTTCTGAGAGGAGTCAAGTGGAGAGCAGGCTCCCGGCAACCAAGCGGAG 473
    |||
Db 512 TTCTGCAAAATTTCTGAGAGGAGTCAAGTGGAGAGCAGGCTCCCGGCAACCAAGCGGAG 571

QY 474 GCGGAGTCCGCGGCGGCGGAGGAGTGTGCGGAGGCGGAGGAGGCGGCGGAGGCGG 533
    |||
Db 572 GCGGAGTCCGCGGCGGCGGAGGAGTGTGCGGAGGCGGAGGAGGCGGCGGAGGCGG 631

QY 534 ACCCGAGAGAGTCAAGGCGGCGGAGGAGTGTGCGGAGGAGGAGGAGGAGGAGGAG 593
    |||
Db 632 ACCCGAGAGAGTCAAGGCGGCGGAGGAGTGTGCGGAGGAGGAGGAGGAGGAGGAG 691

QY 594 CCGGCTCCGGG 604
    |||
Db 692 CCGGCTCCGGG 702
```

RESULT 39
BX381775 931 bp mRNA linear EST 28-APR-2004
LOCUS BX381775

DEFINITION BX381775 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0D1069YH17 5-PRIME, mRNA sequence.
ACCESSION BX381775
VERSION BX381775.2 GI:46833544
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 931)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30458991.
Contact: Genoscope

Genoscope - Centre National de Sequenage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna78-CS0D1069CD09QPic-9074.r>.

FEATURES
SOURCE

```
1. 931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1069YH17"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

```
Query Match 47.9%; Score 435; DB 5; Length 931;
Best Local Similarity 100.0%; Pred. No. 8.2e-210;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CGGAGTGGGCGGCGGCGGAGGAGTGTGCGGAGGCGGCGGAGGCGGCGGAGGCGG 534
    |||
Db 177 CGGAGTGGGCGGCGGCGGAGGAGTGTGCGGAGGCGGCGGAGGCGGCGGAGGCGG 236

QY 535 CCCGAGGAGAGTCAAGAGCCCGGAGAGCTTCTCTGAAGGCAAGTGAACCTGTGACATC 594
    |||
Db 237 CCCGAGGAGAGTCAAGAGCCCGGAGAGCTTCTCTGAAGGCAAGTGAACCTGTGACATC 296

QY 595 CGGCTCCGGGTTGAGAGAGTACTGCGAGCATGAGGCGGAGGCTTGGAGAGGCGGAG 654
    |||
Db 297 CGGCTCCGGGTTGAGAGAGTACTGCGAGCATGAGGCGGAGGCTTGGAGAGGCGGAG 356

QY 655 TCCGCGGCGGCGGCGGCGGAGTGTGCGGAGTGAACGTTTGGGAGGCGGAGGAGT 714
    |||
Db 357 TCCGCGGCGGCGGCGGCGGAGTGTGCGGAGTGAACGTTTGGGAGGCGGAGGAGT 416

QY 715 CTGCGCTCAAGAGAGCTGAGGCTGTGTGTTGTGATCAATCAAGTTCTAGAGCTCTCAT 774
    |||
Db 417 CTGCGCTCAAGAGAGCTGAGGCTGTGTGTTGTGATCAATCAAGTTCTAGAGCTCTCAT 476

QY 775 CTGAGCGCTTTGAGGCGGAGTCAAGTGTGAGGCGGCGGAGGCGGCGGAGGCGG 834
    |||
Db 477 CTGAGCGCTTTGAGGCGGAGTCAAGTGTGAGGCGGCGGAGGCGGCGGAGGCGG 536

QY 835 TTCTGAGTGAAGGCGGCGGAGAGGCTGTGCGGCGGAGGAGGAGGAGGAGGAGT 894
    |||
Db 537 TTCTGAGTGAAGGCGGCGGAGAGGCTGTGCGGCGGAGGAGGAGGAGGAGGAGT 596

QY 895 GTGATGAGGCTGAC 909
```

```
RESULT 40
BP211545
LOCUS
DEFINITION BP211545 Sugano cDNA library, cerebrum Homo sapiens cDNA clone
ACCESSION CBR01216, mRNA sequence.
BP211545
VERSION BP211545.1 GI:52084436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 580)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..580
FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CBR01216"
/cisue_type="cerebrum"
/clone_lib="Sugano cDNA library, cerebrum"
ORIGIN
Query Match 47.5%; Score 432; DB 5; Length 580;
Best Local Similarity 99.8%; Pred. No. 2.8e-208;
Matches 482; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGGATCCGGGTGACCGCGCGGCTGGGAGAGATGAGTGCCTGACTAC 60
DB 98 ATGGCGGATCCGGGTGACCGCGCGGCTGGGAGAGATGAGTGCCTGACTAC 157
QY 61 TACGGATGCTGTGCTTCAACCGTATGTCAGAGTGTGGCGGCACTGACCGAGTGC 120
DB 158 TAGGGATGCTGTGCTTCAACCGTATGTCAGAGTGTGGCGGCACTGACCGAGTGC 217
QY 121 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAGCC 180
DB 218 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAGCC 277
QY 181 CGGGCGCGGAGGCGGCTAGAGCTCTGCTGAGTGGAGCGCGGAGGAGTGGGGGAG 240
DB 278 CGGGCGCGGAGGCGGCTAGAGCTCTGCTGAGTGGAGCGCGGAGGAGTGGGGGAG 337
QY 241 AGCAACCTGGGCTGCTGGGCAACTCTGCGCGTGTGGCGCGCAAGACTTGCTCG 300
DB 338 AGCAACCTGGGCTGCTGGGCAACTCTGCGCGTGTGGCGCGCAAGACTTGCTCG 397
QY 301 CACTGGCGCGCAAGCGGCGCGGCTGCTCAAGAGCTATAGCTATGGCACTTCC 360
DB 398 CACTGGCGCGCAAGCGGCGCGGCTGCTCAAGAGCTATAGCTATGGCACTTCC 457
QY 361 AGCTCTCAAGAGAGAGAGGAGTGTGCGCGTGTGGCGCGCAAGAGAGTTTCA 420
DB 458 AGCTCTCAAGAGAGAGAGGAGTGTGCGCGTGTGGCGCGCAAGAGAGTTTCA 517
QY 421 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGACGCGGAGT 480
DB 518 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGACGCGGAGT 577
```

```
RESULT 41
BP305292
LOCUS
DEFINITION BP305292 Sugano cDNA library, macrophage Homo sapiens cDNA clone
ACCESSION MPG04145, mRNA sequence.
BP305292
VERSION BP305292.1 GI:52234252
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1..583
FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MPG04145"
/cell_type="macrophage"
/clone_lib="Sugano cDNA library, macrophage"
ORIGIN
Query Match 47.2%; Score 429; DB 5; Length 583;
Best Local Similarity 100.0%; Pred. No. 9.3e-207;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 328 GTGTCTCCAGAACGCTTACTATGACACTCCAGCTCTTCAAGAGACAGAGGTAGC 387
DB 59 GTGTCTCCAGAACGCTTACTATGACACTCCAGCTCTTCAAGAGACAGAGGTAGC 118
QY 388 TGCCGTGCGCGTGGCGAGTCAAGAGTTTCTGAATTTCTGACGAGGCTCAGTGGAGACA 447
DB 119 TGCCGTGCGCGTGGCGAGTCAAGAGTTTCTGAATTTCTGACGAGGCTCAGTGGAGACA 178
QY 448 GGTCTCCCTCCCAACCAAGCGGAGCGGAGGTGGGGCGGCGCCAGTGTGTGCCAGA 507
DB 179 GGTCTCCCTCCCAACCAAGCGGAGCGGAGGTGGGGCGGCGCCAGTGTGTGCCAGA 238
QY 508 CGGGCGGAGAGGGGGCCGACCGGACCCAGAGAGTCAAGAGCCCGGACGACTTCC 567
DB 239 CGGGCGGAGAGGGGGCCGACCGGACCCAGAGAGTCAAGAGCCCGGACGACTTCC 298
QY 568 TCTGAAGCAAAAGTGAAGTGAATCGGCTCCGGGTTGAGGAGAGTACTGAGACAT 627
DB 299 TCTGAAGCAAAAGTGAAGTGAATCGGCTCCGGGTTGAGGAGAGTACTGAGACAT 358
QY 628 GGGCGACCTTTGAGAGAGGGCGTGGATCCCGGCGGCGCCAGCGGCTGGCGGCACTG 687
DB 359 GGGCGACCTTTGAGAGAGGGCGTGGATCCCGGCGGCGCCAGCGGCTGGCGGCACTG 418
QY 688 GAGCTTTTGGGCGAGGCGCACGAGTGTGCGTCAAGGAGACTTGGGCTCTGTGTTGT 747
DB 419 GAGCTTTTGGGCGAGGCGCACGAGTGTGCGTCAAGGAGACTTGGGCTCTGTGTTGT 478
QY 748 GACATCAAG 756
DB 479 GACATCAAG 487
```

RESULT 42
CB990003
LOCUS
DEFINITION CB990003 817 bp mRNA linear EST 01-MAY-2003
IMAGE:30347629 5', mRNA sequence.
ACCESSION
CB990003
VERSION
CB990003.1 GI:30284523
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 817)
NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hanson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: NDAM389 row: a column: 14
High quality sequence stop: 566.
Location/Qualifiers
1..817
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30347629"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_id="NIH MGC 147"
/note="Organ: Placenta; Vector: pBluescriptR; Site 1:
ali-XhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTT-3', size-selected for average
insert size 2.3 kb and normalized for ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH-MGC library."

ORIGIN
Query Match 47.0%; Score 427; DB 6; Length 817;
Best Local Similarity 99.8%; Pred. No. 9.6e-206; Indels 0; Gaps 0;
Matches 477; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCGCTATCCGGGTGACCCCGGCGCTGTGGAGAGAGATGCTGACTTAC 60
DB 151 ATGCGCTATCCGGGTGACCCCGGCGCTGTGGAGAGAGATGCTGACTTAC 210
QY 61 TACGGGATGCTGTGCTTACACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 120
DB 211 TACGGGATGCTGTGCTTACACCGTATGTTGAGAGTGTGGCGGGCAACTGACGAGTGC 270
QY 121 GAGCTGAGCTCTGGGCTTTCTGTGTGATGATGAGGCTCTTGGCGCGCGGAGGCTTACGCC 180
DB 271 GAGCTGAGCTCTGGGCTTTCTGTGTGATGATGAGGCTCTTGGCGCGCGGAGGCTTACGCC 330
QY 181 CGGCGCCGACGGCGCTAGAGCTCTCTGTGAGCTGAGCGCCGCGGGCAGTGTGGCGAG 240
DB 331 CGGCGCCGACGGCGCTAGAGCTCTCTGTGAGCTGAGCGCCGCGGGCAGTGTGGCGAG 390
QY 241 AGCAACTGCGGCTGTGTGGGCAACTCTGTGCGGTGTGGCGCCGACGACCTGTGCGG 300
DB 391 AGCAACTGCGGCTGTGTGGGCAACTCTGTGCGGTGTGGCGCCGACGACCTGTGCGG 450

QY 301 CACCTGGCGCGCAAGCGGCGCGGCACTGTCTCCAGAACGCTATAGCTATGACCTCC 360
DB 451 CACCTGGCGCGCAAGCGGCGCGGCACTGTCTCCAGAACGCTATAGCTATGACCTCC 510
QY 361 AGCTCTTCAAGAGAGACAGAGGTTAGTCCGTGCGCGGTGCGGACGTCAAGAGTTTGGCA 420
DB 511 AGCTCTTCAAGAGAGACAGAGGTTAGTCCGTGCGCGGTGCGGACGTCAAGAGTTTGGCA 570
QY 421 AATTCTGACAGGCTGATGAGAGACAGAGCTCCCCCAACCAAGCGGACGCGGCA 478
DB 571 AATTCTGACAGGCTGATGAGAGACAGAGCTCCCCCAACCAAGCGGACGCGGCA 628

RESULT 43
BX458265
LOCUS
DEFINITION BX458265 891 bp mRNA linear EST 06-MAY-2004
5-PRIME, mRNA sequence.
ACCESSION
BX458265
VERSION
BX458265.2 GI:47064757
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 891)
Li W.B., Gruber C., Jesssee, J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 22, 2003 this sequence version replaced gi:31028969.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-Oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 9074.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?ts=CS0DE01AG06QPl&c=9074.r>.
Location/Qualifiers
1..891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE01YM11"
/tissue_type="PLACENTA"
/clone_id="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-Oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 46.9%; Score 426; DB 5; Length 891;
Best Local Similarity 99.8%; Pred. No. 3.1e-205; Indels 1; Gaps 1;
Matches 546; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 328 GTGCTCCAGAAAGCTTACTATAGACCTCCAGCTCTTCAAGAGAGACAGAGGTTAGC 387
DB 111 GTGCTCCAGAAAGCTTACTATAGACCTCCAGCTCTTCAAGAGAGAGAGGTTAGC 170
QY 388 TGCCGTGCGCGTGGCGAGTCAAGAGATTCTCAAAATTTCTAGCAGGCTCAGTGGAGACA 447
DB 171 TGCCGTGCGCGTGGCGAGTCAAGAGATTCTCAAAATTTCTAGCAGGCTCAGTGGAGACA 230
QY 448 GGTCTCCCCCAACCAAGCGGCGAGCGG- CGAGTGTGGGCGCGGCCAGTGTGTGTCAG 506
DB 231 GGTCTCCCCCAACCAAGCGGCGAGCGGAGTGTGGGCGCGGCCAGTGTGTGTCAG 290

QY	507	ACGGCGGGGAGAGGGGCCCCAGCCGACCCGACGAGAGTCAGAGCCCGCAGACTTC	566
Db	291	ACGGCGGGGAGAGGGGCCCCAGCCGACCCGACGAGAGTCAGAGCCCGCAGACTTC	350
QY	567	CTCTGAAGGCAAAAGTGAACCTGTGACATCCGAGCTCCGGGTTTGAGACAGAGTACTGCGAGCA	626
Db	351	CTCTGAAGGCAAAAGTGAACCTGTGACATCCGAGCTCCGGGTTTGAGACAGAGTACTGCGAGCA	410
QY	627	TGGGCGAGCCTTTGAGAGCAGGGGCTGGGCATCCCGGGGGCCCCAGGGGCTGGGCGCGCAGCT	686
Db	411	TGGGCGAGCCTTTGAGAGCAGGGGCTGGGCATCCCGGGGGCCCCAGGGGCTGGGCGCGCAGCT	470
QY	687	GGAGCGTGTTTGGGCGAGGCGCACCGCAGTGTGGGCTCAAGGGAGCTGGGCTCTGTGGTTTG	746
Db	471	GGAGCGTGTTTGGGCGAGGCGCACCGCAGTGTGGGCTCAAGGGAGCTGGGCTCTGTGGTTTG	530
QY	747	TGACATCAAGTCTTCAGAGCTCTCCTATCTGGACGCTCTGGGGGAGACTAACCTGAGTGG	806
Db	531	TGACATCAAGTCTTCAGAGCTCTCCTATCTGGACGCTCTGGGGGAGACTAACCTGAGTGG	590
QY	807	CGCCCTGCTGAGAGCCCTTGCGGGGCGTGTTCCTGACTGAGGCCCTTGCAGAGGCTGTGGG	866
Db	591	CGCCCTGCTGAGAGCCCTTGCGGGGCGTGTTCCTGACTGAGGCCCTTGCAGAGGCTGTGGG	650
QY	867	CCGGGAG 873	
Db	651	CCGGGAG 657	

RESULT	44
BUS07654	
LOCUS	938 bp mRNA linear EST 12-SEP-2002
DEFINITION	AGENCOURT_10095848 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:65016221
ACCESSION	BUS07654
VERSION	BUS07654
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Mumaiyoti; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
AUTHORS	Eumaliya; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	1 (bases 1 to 938)
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Stransberg, Ph.D. Email: csqabbs@mail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LHAM14056 row: f column: 22 High quality sequence stop: 541. Location/Qualifiers 1..938
FEATURES	
source	

```

FEATURES
SOURCE
Location/Qualifiers
1. .938
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6501621"
/issue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NIH MGC 71"
/note="Organ: uterus; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 2.1 kb. "
```

Query Match	Score	DB 5;	Length
Best Local Similarity	Pred. No.	1.2e-200;	
45.9%;	417;		938;
99.8%;			

	Matches	467;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	ATGCGCTATTCGGSGTGCACCCCGGCCCCGCTGCTGGAGAGATGATGCTTGACATAC	60							
Db	65	ATGCGCTATTCGGSGTGCACCCCGGCCCCGCTGCTGGAGAGATGATGCTTGACATAC	124							
QY	61	TACGGGATGCTGTGTGCTTACCGTATGTTGAGGTGTGTGGCGGCACTAACCGAGTGC	120							
Db	125	TACGGGATGCTGTGTGCTTACCGTATGTTGAGGTGTGTGGCGGCACTAACCGAGTGC	184							
QY	121	GAGCTGGAGTCTCTGGGCTTTTCTGCTGATGATAGGCTCTTGCGCGCGCCGAGGCTTTAGCC	180							
Db	185	GAGCTGGAGTCTCTGGGCTTTTCTGCTGATGATAGGCTCTTGCGCGCGCCGAGGCTTTAGCC	244							
QY	181	CGGGCCCGGACGGGCTTAGAGCTCTGTGGAGGTGGAGCGCGCGGGCACTGGCGGAG	240							
Db	245	CGGGCCCGGACGGGCTTAGAGCTCTGTGGAGGTGGAGCGCGCGGGCACTGGCGGAG	304							
QY	241	AGCAACCTGGCGGCTGTGGGGCAACTCTGTGGCGGCTGGCGGCGACGACTGTGTCG	300							
Db	305	AGCAACCTGGCGGCTGTGGGGCAACTCTGTGGCGGCTGGCGGCGACGACTGTGTCG	364							
QY	301	CACCTGGGCGGCAAGCGGCGCGGCGCAGTGTCTCGAAGACGCTATAGCTTAGGCACTCC	360							
Db	365	CACCTGGGCGGCAAGCGGCGCGGCGCAGTGTCTCGAAGACGCTATAGCTTAGGCACTCC	424							
QY	361	AGCTCTTCAAGAGAGACAGAGGTAGCTGTGCTGCCTCGGACGATCAAGCAGTTCTGCA	420							
Db	425	AGCTCTTCAAGAGAGACAGAGGTAGCTGTGCTGCCTCGGACGATCAAGCAGTTCTGCA	484							

```

RESULT 45
LOCUS      BQ947706
DEFINITION BQ947706 1151 bp mRNA linear EST 21-AUG-2002
            AGENCOURT_8931695 NIH_MGC_40 Homo sapiens CDNA clone IMAGE:6484004
            5', mRNA sequence.
ACCESSION  BQ947706
VERSION     BQ947706.1  GI:22363184
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 1151)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
JOURNAL     Tissue Procurement: DCPD/DTP
COMMENT      CDNA Library Preparation: Robin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            Plate: L1CM2669 row: h column: 21
            High quality sequence stop: 410.
FEATURES
            Location/Qualifiers
                1..1151

```

```

/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NH_MGC_40"
/notes="Organ: prostate; Vector: pOT7; Site_1: XhoI,
Site_2: EcoRI; cDNA made by oligo-dT priming.

```


Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 45.7%; Score 415; DB 5; Length 1151;
Best Local Similarity 100.0%; Pred. No. 1,2e-199;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 466 CGGACGCGCGAGTCGGGGCCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGGGCC 525
Db 1 CGGACGCGCGAGTCGGGGCCGCGCCAGTGTGTGTCAGACGCGCGGAGAGGGGCC 60
QY 526 CCACCGCGACCCCGACAGACATGACAGCCCGCCAGACCTTCTCTGAAGGCAAGTGCAC 585
Db 61 CCACCGCGACCCCGACAGACATGACAGCCCGCCAGACCTTCTCTGAAGGCAAGTGCAC 120
QY 586 TGTGACATCCGGCTCCGGGTTGAGACAGACTGTCAGACATGGGCGACGCTTGAGAGAG 645
Db 121 TGTGACATCCGGCTCCGGGTTGAGACAGACTGTCAGACATGGGCGACGCTTGAGAGAG 180
QY 646 GGCCTGGCATCCCGCGGCGCCGAGCGCTGGCGCGGACGTGAAGTGTGGCGAGGCC 705
Db 181 GGCCTGGCATCCCGCGGCGCCGAGCGCTGGCGCGGACGTGAAGTGTGGCGAGGCC 240
QY 706 ACCGACATGTCGCTGCAAGGAGACTGGGCTGTGTGTTGTGACATGAAGTTCTCAGAG 765
Db 241 ACCGACATGTCGCTGCAAGGAGACTGGGCTGTGTGTTGTGACATGAAGTTCTCAGAG 300
QY 766 CTCTCTATCTGAGAGCGCTTCTGGGGCGAGTACCTGATGGGCGCCCTGTCAGAGCCCTG 825
Db 301 CTCTCTATCTGAGAGCGCTTCTGGGGCGAGTACCTGATGGGCGCCCTGTCAGAGCCCTG 360
QY 826 CGGGCGGTGTCCTGACTGAGAGCGCTGCGAGAGCGTGTGGCGCGGAGCGTGTTC 880
Db 361 CGGGCGGTGTCCTGACTGAGAGCGCTGCGAGAGCGTGTGGCGCGGAGCGTGTTC 415

RESULT 46
LOCUS B1769587 837 bp mRNA linear EST 25-SEP-2001
DEFINITION 603054964P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204607 5',
mRNA sequence.

ACCESSION B1769587
VERSION B1769587.1 GI:15761165
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 837)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: LAM1512 row: p column: 16

High quality sequence stop: 830.

Location/Qualifiers

1. 837
/Organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:5204607"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SportE;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source:
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 45.1%; Score 410; DB 4; Length 837;
Best Local Similarity 99.8%; Pred. No. 4.4e-197;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTGACCCCGCGCGTGTGAGAGAGATGATGCTGACTAC 60
Db 40 ATGGCGCTATCCGGGTGACCCCGCGCGTGTGAGAGAGATGATGCTGACTAC 99
QY 61 TACGGATGCTGTGCTTCAACCGTATGTGAGGTGTGGCGGCAACTGACCGAGTGC 120
Db 100 TACGGATGCTGTGCTTCAACCGTATGTGAGGTGTGGCGGCAACTGACCGAGTGC 159
QY 121 GAGCTGAGCTCCCTGGGCTTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAGCC 180
Db 160 GAGCTGAGCTCCCTGGGCTTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAGCC 219
QY 181 CGGGCCCGACAGCGGCTAGAGCTCTGCTGAGAGTGAAGCGCCGCGGCAAGTGGCGAG 240
Db 220 CGGGCCCGACAGCGGCTAGAGCTCTGCTGAGAGTGAAGCGCCGCGGCAAGTGGCGAG 279
QY 241 AGCAACTGCGGCTGTGGGCAACTCTGTGCGCTGTGGCGCGCAAGTCTGCTCCG 300
Db 280 AGCAACTGCGGCTGTGGGCAACTCTGTGCGCTGTGGCGCGCAAGTCTGCTCCG 339
QY 301 CACCTGGCGGCAAGCGGCGCGGCGAGTGTCTTCAAGAGCTTACTTATGACACTCC 360
Db 340 CACCTGGCGGCAAGCGGCGCGGCGAGTGTCTTCAAGAGCTTACTTATGACACTCC 399
QY 361 AGCTCTCAAGAGAGAGAGGAGTGTGCGGCTGCGGCGGCAAGTGTGCA 420
Db 400 AGCTCTCAAGAGAGAGAGGAGTGTGCGGCTGCGGCGGCAAGTGTGCA 459
QY 421 AATTCTCAGAGAGGTCAAGTGTGAGAGAGGCTCCCGCCCAAC 461
Db 460 AATTCTCAGAGAGGTCAAGTGTGAGAGAGGCTCCCGCCCAAC 500

RESULT 47
LOCUS BM547754

DEFINITION AGENCOURT 6507032 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5727891
5', mRNA sequence.

ACCESSION BM547754

VERSION BM547754.1 GI:18781793

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1233)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
 Plate: LLM2722 row: 1 column: 04
 High quality sequence stop: 542.

FEATURES

source

1..1233

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAG:5727891"

/tissue_type="hippocampus"

/lab_host="DH10B"

/clone_lib="NIH MGC 124"

/note="Organ: brain; Vector: pCMV-Sport6; Site 1: EcoRV
 (destroyed); Site 2: NotI; RNA source: male hippocampus,
 age 27. Library is oligo-dT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.4 kb, insert size range 0.9-4 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 012."

ORIGIN

Query Match 44.8%; Score 407; DB 4; Length 1233;
 Best Local Similarity 99.8%; Pred. No. 1.5e-195;
 Matches 457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCGCTATCCGGGTGAGACCCCGCCGTCGGAGAGAGATGAGTCCCTGACTAC 60
 73 ATGGCGCTATCCGGGTGAGACCCCGCCGTCGGAGAGAGATGAGTCCCTGACTAC 120
 61 TACGGAGATGCTGCTTCAACCGATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
 133 TACGGAGATGCTGCTTCAACCGATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 192
 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 180
 193 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 252
 181 CGGGCCCGCAGCGGCGCTAGAGCTCTGCTGAGTGGAGCGCGCGGCGAGTGGCGGAG 240
 253 CGGGCCCGCAGCGGCGCTAGAGCTCTGCTGAGTGGAGCGCGCGGCGAGTGGCGGAG 312
 241 AGCAACTGCGGCTGCTGGGCAACTCTGCGCTGCTGCGCCGACGACTGCTGCCG 300
 313 AGCAACTGCGGCTGCTGGGCAACTCTGCGCTGCTGCGCCGACGACTGCTGCCG 372
 301 CACCTGCGCGCAAGCGGCGCGGCGCAAGTCTTCAGAACTATAGCTATGCACTCC 360
 373 CACCTGCGCGCAAGCGGCGCGGCGCAAGTCTTCAGAACTATAGCTATGCACTCC 432
 361 AGCTCTTCAAGAAGAGAGAGAGGAGTACGCTGCGCGCGGCGAGTAAAGATTTCGA 420
 433 AGCTCTTCAAGAAGAGAGAGGAGTACGCTGCGCGCGGCGAGTAAAGATTTCGA 492
 421 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCCC 458
 493 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCCC 530

RESULT 48
 BP253687

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

582 bp mRNA linear EST 15-SEP-2004

BP253687 Sugano cDNA library, kidney epithelial cell Homo sapiens

BP253687

BP253687.1 GI:52135968

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 582)

AUTHORS

Suzuki,Y., Yamashita,R., Shiroya,M., Sakakibara,Y., Chiba,J.,

Mizushima-Sugano,J., Nakai,K. and Sugano,S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

Genome Res. 14 (9), 1711-1718 (2004)

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1..582

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HRC04630"

/tissue_type="kidney"

/cell_type="epithelial cell"

/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN

Query Match 44.3%; Score 403; DB 5; Length 582;
 Best Local Similarity 99.8%; Pred. No. 1.6e-193;
 Matches 453; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGGCGCTATCCGGGTGAGACCCCGCCGTCGGAGAGAGATGAGTCCCTGACTAC 60
 120 ATGGCGCTATCCGGGTGAGACCCCGCCGTCGGAGAGAGATGAGTCCCTGACTAC 179
 61 TACGGAGATGCTGCTTCAACCGATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 120
 180 TACGGAGATGCTGCTTCAACCGATGTTGAGAGTGGGCGGCAACTGACCGAGTGC 239
 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 180
 240 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAGCC 239
 181 CGGGCCCGCAGCGGCGCTAGAGCTCTGCTGAGTGGAGCGCGCGGCGAGTGGCGGAG 240
 300 CGGGCCCGCAGCGGCGCTAGAGCTCTGCTGAGTGGAGCGCGCGGCGAGTGGCGGAG 359
 241 AGCAACTGCGGCTGCTGGGCAACTCTGCGCTGCTGCGCCGACGACTGCTGCCG 300
 360 AGCAACTGCGGCTGCTGGGCAACTCTGCGCTGCTGCGCCGACGACTGCTGCCG 419
 301 CACCTGCGCGCAAGCGGCGCGGCGCAAGTCTTCAGAACTATAGCTATGCACTCC 360
 420 CACCTGCGCGCAAGCGGCGCGGCGCAAGTCTTCAGAACTATAGCTATGCACTCC 479
 361 AGCTCTTCAAGAAGAGAGAGGAGTACGCTGCGCGCGGCGAGTAAAGATTTCGA 420
 480 AGCTCTTCAAGAAGAGAGGAGTACGCTGCGCGCGGCGAGTAAAGATTTCGA 539
 421 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCC 454
 540 AATTCTCAGCAGGCTCAGTGGAGACAGGCTCCCC 573

RESULT 49

CN373296

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

402 bp mRNA linear EST 16-MAY-2004

CN373296 170059942874 GRN_PRENED Homo sapiens cDNA 5', mRNA sequence.

CN373296

CN373296.1 GI:47373230

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Brandenberger,R., Wei,H., Zhang,S., Lei,S., Muraige,J., Fisk,G.J.,

TITLE
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL
Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT
Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 402 Std Error: 0.00.
Location/Qualifiers
1. 402

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN PRENU"
/note="Oligo dt primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."

ORIGIN

Query Match 43.2%; Score 393; DB 7; Length 402;
Best Local Similarity 100.0%; Pred. No. 2e-186;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 GTGTCTCAGAAACCTATAGTATGACACCTCTTCAAGAGAGACAGAGTAC 387
Db 10 GTGTCTCAGAAACCTATAGTATGACACCTCTTCAAGAGAGAGAGTAC 69
Qy 388 TGCCGTGCGGTGCGAGTCAAGAGATTCTGCAAAATTTCTGACAGAGGTCACTGGAGACA 447
Db 70 TGCCGTGCGGTGCGAGTCAAGAGATTCTGCAAAATTTCTGACAGAGGTCACTGGAGACA 129
Qy 448 GGGTCCCCCCCAACAGCGGCGAGGTGGGGCCCGCCCACTGCTGTGCGACA 507
Db 130 GGGTCCCCCCCAACAGCGGCGAGGTGGGGCCCGCCCACTGCTGTGCGACA 189
Qy 508 CGGCGGCGAGAGGGGCCCGACCGCACCCAGAGAGAGTCAAGAGCCCGACACTTCC 567
Db 190 CGGCGGCGAGAGGGGCCCGACCGCACCCAGAGAGAGTCAAGAGCCCGACACTTCC 249
Qy 568 TCTGAAGGCAAGTACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGCGACAT 627
Db 250 TCTGAAGGCAAGTACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGCGACAT 309
Qy 628 GGGCGACCTTTGAGAGAGGGGCTGGCATCCGGGGGCCCGGCGGCTGGCGGCACTG 687
Db 310 GGGCGACCTTTGAGAGAGGGGCTGGCATCCGGGGGCCCGGCGGCTGGCGGCACTG 369
Qy 688 GACGTGTTGGGCGAGCCGACCGACAGTGTGCGC 720
Db 370 GACGTGTTGGGCGAGCCGACCGACAGTGTGCGC 402

RESULT 50
AM245758 475 bp mRNA linear EST 07-JAN-2000
LOCUS 28222785.Sprime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822785 5',
DEFINITION mRNA sequence.
ACCESSION AM245758
VERSION AM245758.1 GI:5588751
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 475)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other_ESTRs: 2822785.Sprime
Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP cDNA Library Preparation: Ling
Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
Consortium (LIND) DNA Sequencing by: Berkeley MGC sequencing
project Clone distribution: MGC clone distribution information can
be found through the I.M.A.G.E. Consortium/LIND at:
www.bio.lind.gov/bdip/image/image.html Base Calling / Quality
Scores: PHRED from University of Washington Genome Center. Vector
Trimming: cross match from University of Washington Genome Center
PHRAP suite. Poly-T identification: patmatch.pl from Berkeley
Drosophila Genome Project. University of Washington Genome Center:
<http://www.genome.washington.edu>
Plate: LUCM10 row: F column: 2
High quality sequence stop: 430.
Location/Qualifiers
1. 475

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2822785"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 42.7%; Score 388; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 6.9e-186;
Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 522 GGCCCCAGCCGACCCAGAGAGTCAAGAGCCCGACACTTCTCTGAAGGCAAGT 581
Db 8 GGCCCCAGCCGACCCAGAGAGTCAAGAGCCCGACACTTCTCTGAAGGCAAGT 67
Qy 582 GACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGCGACATGGGCGACGCTTGA 641
Db 68 GACCTGTGACATCCGGCTCCGGGTTGAGAGAGTACTGCGACATGGGCGACGCTTGA 127
Qy 642 GCAAGGCGTGCATCCCGGGGCCCGGCGGCTGGCGGCGGCACTGAGTGTGGGCA 701
Db 128 GCAAGGCGTGCATCCCGGGGCCCGGCGGCTGGCGGCGGCACTGAGTGTGGGCA 187
Qy 702 GGCACCGAGTGTGGCTCAAGGAGACTGGGCTCTGTGTTTGTGATCAAGTCTC 761
Db 188 GGCACCGAGTGTGGCTCAAGGAGACTGGGCTCTGTGTTTGTGATCAAGTCTC 247
Qy 762 AGAGCTTCTCTATCTGAGACGCTTCTGGGGGCACTACTAGTGGCGCCCTGTGACGC 821
Db 248 AGAGCTTCTCTATCTGAGACGCTTCTGGGGGCACTACTAGTGGCGCCCTGTGACGC 307
Qy 822 CTCGCGGGGGGTGTTCTCTGACTGAGGCGCTGCGAGAGGCTGTGGCGGAGAGTCTTCG 881
Db 308 CTCGCGGGGGGTGTTCTCTGACTGAGGCGCTGCGAGAGGCTGTGGCGGAGAGTCTTCG 367
Qy 882 CTCGCTGTGAGTGTGATGAGGCTGAC 909
Db 368 CTCGCTGTGAGTGTGATGAGGCTGAC 395

Search completed: March 24, 2005, 03:06:24

Thu Mar 24 07:06:31 2005

us-10-030-271-1.oligo.rst

Page 32

Job time : 3270.01 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 16:09:50 ; Search time 526.777 Seconds
(without alignments)
10215.030 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctatccggtcgcac.....tcagtgtgatgagctgac 909

Scoring table: OLIGO_NUC

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

- 1: N_Geneseq_16Dec04:*
- 2: geneseqn1980s:*
- 3: geneseqn1990s:*
- 4: geneseqn2000s:*
- 5: geneseqn2001as:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	909	100.0	909	5	AAf27407 Human apo
2	909	100.0	1883	5	AAf27408 Human apo
3	858	94.4	981	10	ADc79259 Human DED
4	858	94.4	1230	6	ABA94362 Human APR
5	858	94.4	1924	6	AAAD40080 Human DED
6	858	94.4	1924	6	AAAD59062 Human DED
7	858	94.4	1979	12	ADQ086891 Human tun
8	858	94.4	2045	3	AAAS3283 Apoptosis
9	858	94.4	2044	3	AAAF18296 lung can
10	768	84.5	2044	3	AAAF18296 lung can
11	538	59.2	1966	6	AAAS62603 CDNA sequ
12	531	58.4	1570	4	AAH99646 Human pro
13	534	36.7	626	4	AAH99646 Human pro
14	281	30.9	1042	4	AAI60747 Human pol
15	281	30.9	1067	5	AAI58961 Human pol
16	281	30.9	1067	5	ADQ09183 DNA encod
17	281	30.9	1067	5	ADQ09183 DNA encod
18	277	30.5	11084	12	ADQ18808 Human sof
19	252	27.7	303	6	AAAD40075 Human DED
20	252	27.7	303	6	AAAD59057 Human DED

21	187	20.6	1084	6	ABL39692 Human NS
22	187	20.6	1106	6	ABL39693 Human NS
23	177	19.5	603	5	AAf93937 Primer sp
24	151	16.6	451	9	ACH41209 Human foe
25	96	10.6	484	9	ACH25056 Human adu
26	25	2.8	25	10	ADc79263 Human DED
27	27	2.4	838	5	ADL44878 Human ova
28	21	2.3	333	5	ADL39650 Human ova
29	20	2.2	21	5	AAf27413 Human apo
30	20	2.2	23	5	AAf27416 Human apo
31	20	2.2	2610	9	ACC59397 Microbial
32	20	2.2	2673	11	ABD04389 Pseudomon
33	20	2.2	2715	11	ABD04082 Pseudomon
34	20	2.2	2814	11	ABD04271 Pseudomon
35	19	2.1	360	10	ADc78561 35 amino
36	19	2.1	579	13	ADQ55265 Novel can
37	19	2.1	606	5	AAH52281 Human AFP
38	19	2.1	943	6	ABK70016 CDNA enco
39	19	2.1	943	9	ADA01379 Human PRO
40	19	2.1	943	9	ADA43808 Human CDN
41	19	2.1	943	9	ADA43576 Human CDN
42	19	2.1	943	9	ADA01251 Human PRO
43	19	2.1	943	9	ADA01135 Human CDN
44	19	2.1	943	9	ADA43692 Human CDN
45	19	2.1	943	9	ADA06954 Human PRO
46	19	2.1	943	9	ADA08442 Novel hum
47	19	2.1	943	9	ADB99735 Human PRO
48	19	2.1	943	9	ADB87018 Human PRO
49	19	2.1	943	9	ADB66173 Human CDN
50	19	2.1	943	10	ADB99851 Human PRO
51	19	2.1	943	10	ADB99506 Novel hum
52	19	2.1	943	10	ADB66057 Human CDN
53	19	2.1	943	10	ADC23455 Human CDN
54	19	2.1	943	10	ADC26148 Human PRO
55	19	2.1	943	10	ADc04975 Human PRO
56	19	2.1	943	10	ADc11281 Human PRO
57	19	2.1	943	10	ADc88212 Human PRO
58	19	2.1	943	10	ADc5507 Human CDN
59	19	2.1	943	10	ADc06437 Human PRO
60	19	2.1	943	10	ADc38212 Human PRO
61	19	2.1	943	10	ADc88328 Human PRO
62	19	2.1	943	10	ADc90909 Human CDN
63	19	2.1	943	10	ADc99464 Human CDN
64	19	2.1	943	10	ADc06557 Human PRO
65	19	2.1	943	10	ADc05508 Human PRO
66	19	2.1	943	10	ADc82509 Human PRO
67	19	2.1	943	12	ADc51762 Human CDN
68	19	2.1	943	12	ADc37736 Human CDN
69	19	2.1	943	12	ADc37736 Human CDN
70	19	2.1	943	12	ADc37620 Human CDN
71	19	2.1	943	12	ADc95391 Human CDN
72	19	2.1	943	12	ADc38091 Human PRO
73	19	2.1	943	12	ADc76180 Human PRO
74	19	2.1	943	12	ADc39503 Human PRO
75	19	2.1	943	12	ADc04307 Human PRO
76	19	2.1	943	12	ADc39904 Human PRO
77	19	2.1	943	12	ADc19769 Human PRO
78	19	2.1	943	12	ADc77347 Human CDN
79	19	2.1	943	12	ADc65455 Human PRO
80	19	2.1	943	12	ADc76064 Human PRO
81	19	2.1	943	12	ADc7975 Human PRO
82	19	2.1	943	12	ADc64585 Human PRO
83	19	2.1	943	12	ADc38990 Human PRO
84	19	2.1	943	12	ADc51994 Human CDN
85	19	2.1	943	12	ADc91025 Human CDN
86	19	2.1	943	12	ADc38804 Human PRO
87	19	2.1	943	12	ADc75504 Human PRO
88	19	2.1	943	12	ADc06320 Human PRO
89	19	2.1	943	12	ADc00180 Human CDN
90	19	2.1	943	12	ADc38668 Human PRO
91	19	2.1	943	12	ADc39619 Human PRO
92	19	2.1	943	12	ADc89224 Human PRO
93	19	2.1	943	12	ADc88991 Human PRO

XX Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
 KM DED; caspase family cleavage domain; pro-apoptotic; drug screening;
 KM cell proliferation; ischaemic disease; chronic viral disease; ss.
 OS Homo sapiens.
 XX MO200104300-A1.
 PN 18-JAN-2001.
 XX 18-JAN-2001.
 PD 06-JUL-2000; 2000WO-JP004516.
 XX PF 08-JUL-1999; 99JP-01194179.
 XX PR 18-OCT-1999; 99US-0159586P.
 XX PA (HELI-) HELIX RES INST.
 PI Ota T, Isozaki T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
 XX WPI: 2001-138348/14.
 DR P-PSDB: AAB60387.
 XX PT Polynucleotide encoding an apoptosis-associated factor protein with death
 PT effector domain and caspase family-cleavage domain, useful in regulating
 PT diseases with cell proliferation.
 XX Claim 3; Page 44-47; 53pp; Japanese.
 XX The invention relates to a novel human apoptosis-associated factor
 CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
 CC effector domain (DED) and a caspase family cleavage domain and is capable
 CC of inducing apoptosis in cells. The invention also relates to nucleic
 CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
 CC (particularly dominant negative variants); vectors and host cells
 CC comprising a nucleic acid which encodes an apoptosis-associated factor
 CC of the invention; the recombinant production of the protein; an antibody
 CC against the protein; and methods of screening for compounds which can
 CC regulate apoptosis. The apoptosis-related factor is useful in regulating
 CC diseases associated with cell proliferation and in screening drug
 CC candidates e.g., for regulating cell proliferation or cell death in
 CC ischaemic diseases and chronic viral diseases. The present sequence
 CC represents a full-length cDNA encoding the human apoptosis-associated
 CC factor NT2RM1000558
 XX
 SQ Sequence 1883 BP; 343 A; 595 C; 588 G; 357 T; 0 U; 0 Other;
 Query Match 100.0%; Score 909; DB 5; Length 1883;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 424 CACCTGGCGCGCAAGCGGCGCGCCAGTGTCTCCAGAACGCTATAGTGCACCTCC 483
 QY AGCTCTTCAAGAGAGACAGAGAGGTAGTGCCTGCGCCGTGGAGTCAAGAGATTCTGCA 420
 DB 361 ||||||| 484 AGCTCTTCAAGAGAGAGAGGTAGTGCCTGCGCCGTGGAGTCAAGAGATTCTGCA 543
 QY 421 AATTCTCAGCAGAGGTCAAGTGGAGACAGAGCTCCCCCAACCAAGCGGACGGCGGAGT 480
 DB 544 AATTCTCAGCAGAGGTCAAGTGGAGACAGAGCTCCCCCAACCAAGCGGACGGCGGAGT 603
 QY 481 CGGGCCCGCCCAAGTGTGTGTCAGACGCGCGGAGAGAGGGCCCCACGCCACCCAG 540
 DB 604 CGGGCCCGCCCAAGTGTGTGTCAGACGCGCGGAGAGAGGGCCCCACGCCACCCAG 663
 QY 541 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAGCAAGAGACCTGTGATCATCCGAGTTC 600
 DB 664 CAGCAGTCAGAGCCCGCAGACCTTCTCTGAAGCAAGAGACCTGTGATCATCCGAGTTC 723
 QY 601 CGGGTTCAGAGCAGATCTGAGCAGATGAGCCAGCCTTGAAGCAGAGCGGTGACATCCCGG 660
 DB 724 CGGGTTCAGAGCAGATCTGAGCAGATGAGCCAGCCTTGAAGCAGAGCGGTGACATCCCGG 783
 QY 661 CGGCCCCAGGCTGTCGCGCGCCAGCTGAGCCTGTTTGGGCAAGCCAGCTGTGCGC 720
 DB 784 CGGCCCCAGGCTGTCGCGCGCCAGCTGAGCCTGTTTGGGCAAGCCAGCTGTGCGC 843
 QY 721 TCAAGGGACCTGGGCTGTGTGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 780
 DB 844 TCAAGGGACCTGGGCTGTGTGTTTGACATCAAGTTCTCAGAGCTCTCTATCTGGAC 903
 QY 781 GCCTTTCGGGGCGACTACTAGTGGCGCCCTGCTGACAGGCCCTGCGGGGCGTTCCTG 840
 DB 904 GCCTTTCGGGGCGACTACTAGTGGCGCCCTGCTGACAGGCCCTGCGGGGCGTTCCTG 963
 QY 841 ACTAGAGCCCTTCGAGAGAGGCTGTGGGCGCGGAGGCTGTTCCTCTGCTGTGATGTGAT 900
 DB 964 ACTAGAGCCCTTCGAGAGAGGCTGTGGGCGCGGAGGCTGTTCCTCTGCTGTGATGTGAT 1023
 QY 901 GAGGCTGAC 909
 DB 1024 GAGGCTGAC 1032

RESULT 3
 ADC79259
 ID ADC79259 standard; cDNA; 981 BP.
 XX
 XX AC ADC79259;
 AC
 DT 01-JAN-2004 (first entry)
 XX
 DE Human DED2 encoding cDNA SEQ ID NO:1.
 XX
 KW human; death effector domains containing DNA-binding protein;
 KW DED-containing DNA-binding protein; DED2; cell death; gene therapy;
 KW cytostatic; cancer; chronic myeloid leukaemia; gene; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..981
 FT CDS /*tag= a
 FT /product= "DED2"
 OS MO2003054195-A1.
 PN 03-JUL-2003.
 PD 20-DEC-2002; 2002WO-JP013371.
 XX PF 20-DEC-2001; 2001JP-00387854.
 XX PR 18-JUL-2002; 2002JP-00209458.
 XX

PA (MORG) MORINAGA MILK IND CO LTD.
 XX
 PI Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
 XX
 DR WPI: 2003-569246/53.
 DR P-PSDB; ADC79260.
 XX
 PT DNA encoding cell death proteins for treatment of kidney, large intestine
 PT and prostate cancers and leukemia.
 XX
 PS Claim 2; Page 18-20; 26pp; Japanese.
 XX
 CC The present sequence encodes a human death effector domains (DED)
 CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that
 CC causes cell death. Also described: (1) primer and probe for investigation
 CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
 CC cytosolic activity. DEDD2 can be used in the diagnosis and treatment of
 CC cancers of the kidney, large intestine and prostate, and acute and
 CC chronic myeloid leukaemia.
 XX
 SQ Sequence 981 BP, 151 A; 303 C; 363 G; 164 T; 0 U; 0 Other;
 Query Match 94.4%; Score 858; DB 10; Length 981;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 60
 DB 1 ATGGCGCTATCCGGGTGACCCCGCCGCTGCTGGAGAGAGATGAGTGCCTGACTAC 60
 QY 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCCGGCACTGACCGAGTGC 120
 DB 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGTGGCCGGCACTGACCGAGTGC 120
 QY 121 GAGCTGAGAGCTCCGCTTCTGCTGATGAGGCTCTGCGCCGCGCGAGGCTTTCAGCC 180
 DB 121 GAGCTGAGAGCTCCGCTTCTGCTGATGAGGCTCTGCGCCGCGCGAGGCTTTCAGCC 180
 QY 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGTGGAGCGCGGGCAGTGGCGAG 240
 DB 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGAGTGGAGCGCGGGCAGTGGCGAG 240
 QY 241 AGCAACTGCTGGCTGTGGGCACTCTGCGCTGCTGGCCCGCCAGACTGCTGCG 300
 DB 241 AGCAACTGCTGGCTGTGGGCACTCTGCGCTGCTGGCCCGCCAGACTGCTGCG 300
 QY 301 CACCTGGCGCGCAGCGGCGCGCGCACTGCTGAGAGCTATGAGTATGAGCACTCC 360
 DB 301 CACCTGGCGCGCAGCGGCGCGCGCACTGCTGAGAGCTATGAGTATGAGCACTCC 360
 QY 361 AGCTCTTCAAGAGAGAGAGAGAGTGGTCCGCTGCGCGCTGCGGCAAGTTCGCA 420
 DB 361 AGCTCTTCAAGAGAGAGAGAGAGTGGTCCGCTGCGCGCTGCGGCAAGTTCGCA 420
 QY 421 AATTCTCAGCAGGCTCAGTGGGAGAGAGGCTCCCCCAACCAAGCGGCGAGGAGT 480
 DB 421 AATTCTCAGCAGGCTCAGTGGGAGAGAGGCTCCCCCAACCAAGCGGCGAGGAGT 480
 QY 481 CGGGGGCGGGCCAGTGGTGTGTCAGACGCGCGGAGAGAGGGGGCCAGCGCCAC 540
 DB 481 CGGGGGCGGGCCAGTGGTGTGTCAGACGCGCGGAGAGAGGGGGCCAGCGCCAC 540
 QY 541 CAGCAGTCAGAGCCCGCAGACTTCTCTGTAAGGCAAGTGAAGTGAATCCGAGTCC 600
 DB 541 CAGCAGTCAGAGCCCGCAGACTTCTCTGTAAGGCAAGTGAAGTGAATCCGAGTCC 600
 QY 601 CGGGTTGAGAGAGTACTGCGAGATGGGCAAGCTTGGAGCGAGCGGCTGTCATCCGG 660
 DB 601 CGGGTTGAGAGAGTACTGCGAGATGGGCAAGCTTGGAGCGAGCGGCTGTCATCCGG 660
 QY 661 CGGGCCCGAGGGCTGGGGCGGCGCACTGAGAGTGTGGGCGAGCGCACCGCAGTCTCGC 720
 DB 661 CGGGCCCGAGGGCTGGGGCGGCGCACTGAGAGTGTGGGCGAGCGCACCGCAGTCTCGC 720

QY 721 TCAGGACCTGGGCTCTGTGTGTTGTGACATGAAGTTCTCAGAGCTCTCTATCTGGAC 780
 DB 721 TCAGGACCTGGGCTCTGTGTGTTGTGACATGAAGTTCTCAGAGCTCTCTATCTGGAC 780
 QY 781 GCCTTCGGGGGACATCAACCTGAGTGGCGGCTGCTGACAGGCGCTGGGGGCGTTCTG 840
 DB 781 GCCTTCGGGGGACATCAACCTGAGTGGCGGCTGCTGACAGGCGCTGGGGGCGTTCTG 840
 QY 841 ACTGAGGCTCTGGAGAGGCTGTGGGCGGAGAGCTGTTGCTGCTGCTGATCAGTGGAT 900
 DB 841 ACTGAGGCTCTGGAGAGGCTGTGGGCGGAGAGCTGTTGCTGCTGCTGATCAGTGGAT 900
 QY 901 GAGGCTGAC 909
 DB 901 GAGGCTGAC 909
 RESULT 4
 ABA94362
 ID ABA94362 standard; DNA; 1230 BP.
 XX
 AC ABA94362;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human APRG polypeptide (Incyte ID. 3102521CD1) encoding cDNA.
 XX
 KW APRG; apoptosis regulator; cytosolic; antiatherosclerotic; osteopathic;
 KW antiarteriosclerotic; hepatotropic; antiparasitic; antihelminthic; human;
 KW antiallergic; antineuritic; antiaschemic; antithyroid; anti-HIV; cancer;
 KW antineoplastic; antidiabetic; antigout; nephrotropic; ophthalmological;
 KW immunosuppressive; dermatological; antitumor; antineuritic; fungicide;
 KW anticholinergic; antibacterial; virucide; antiparasitic; protozoacide;
 KW tranquilizer; vulnery; gynecological; vasotropic; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 127..1107
 FT /*tag= a
 FT /product= "APRG polypeptide"
 FT
 PN W0200192527-A2.
 XX
 PD 06-DEC-2001.
 XX
 PE 30-MAY-2001; 2001WO-US017581.
 XX
 PR 01-JUN-2000; 2000US-0209407P.
 PR 30-NOV-2000; 2000US-0250326P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Tang YT, Azimzai Y, Yue H, Burford N, Ding L, Elliott VS;
 PI Paterson C, Baughn MK;
 XX
 DR WPI: 2002-114350/15.
 DR P-PSDB; ABB07263.
 XX
 PT Novel human apoptosis regulator polypeptides and polynucleotides for
 PT diagnosing, preventing, treating cell proliferative, immunological use,
 PT reproductive disorders and for identifying modulators of therapeutic use.
 XX
 PS Claim 5; Page 101; 103pp; English.
 XX
 CC The invention provides human apoptosis regulator (APRG) polypeptides and
 CC polynucleotides. The APRG polypeptides, polynucleotides and modulators
 CC are useful for diagnosis, treatment and prevention of cell proliferative,
 CC immunological and reproductive disorders. The cell proliferative
 CC disorders include cancers, actinic keratosis, arteriosclerosis,
 CC atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, and
 CC immunological disorders include acquired immunodeficiency syndrome

CC (AIDS), adult respiratory distress syndrome, Addison's disease,
 CC ankylosing spondylitis, amyloidosis, allergies, anemia, osteoporosis,
 CC autoimmune hemolytic anemia, asthma, autoimmune thyroiditis, Crohn's
 CC disease, contact dermatitis, diabetes mellitus, gout, Graves' disease,
 CC glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
 CC erythematosus, systemic sclerosis, ulcerative colitis, hemodialysis,
 CC urethritis, viral, bacterial, fungal, parasitic, protozoal, helminthic
 CC infections and trauma. Reproductive disorders include disorders of
 CC prolactin production, infertility, endometriosis, polycystic ovary
 CC syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,
 CC disruptions of spermatogenesis, cancer of testis and prostate, impotence,
 CC carcinoma of male breast and gynecomastia. The APRG polynucleotides are
 CC useful for creating knockin humanized animals or transgenic animals to
 CC model human disease and to detect and quantify gene expression in
 CC biopsied tissues in which expression of APRG is correlated with disease.
 CC APRG, fragments of it and antibodies specific for APRG are useful as
 CC elements on a microarray which is useful to monitor or measure protein-
 CC protein interactions, drug-target interactions and gene expression
 CC profiles. The present sequence represents a human APRG polypeptide
 CC encoding cDNA

XX Sequence 1230 BP; 205 A; 382 C; 433 G; 210 T; 0 U; 0 Other;

XX Query Match 94.4%; Score 858; DB 6; Length 1230;
 XX Best Local Similarity 99.9%; Pred. No. 0;
 XX Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGTATCCGGGTGAGACCCGGCCGCTGGGAGAGATGAGTCTGAGATAC 60
 DB 127 ATGGCGTATCCGGGTGAGACCCGGCCGCTGGGAGAGATGAGTCTGAGATAC 186
 QY 61 TACGGAGTGTGTGCTTCAACCTATGTTCAAGGTGTGGCGGGCAATGACGAGTGC 120
 DB 187 TAGGGATGTGTGTGCTTCAACCTATGTTCAAGGTGTGGCGGGCAATGACGAGTGC 246
 QY 121 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCCGGAGGCTTAC 180
 DB 247 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCCGGAGGCTTAC 306
 QY 181 CGGGCCCGGAGCGGCGCTTGAAGCTCTGCTGAGTGAAGCGCGCGGAGTGGGAG 240
 DB 307 CGGGCCCGGAGCGGCGCTTGAAGCTCTGCTGAGTGAAGCGCGCGGAGTGGGAG 366
 QY 241 AGCAACCTGGCGCTGCTGGGCAACTCTGCGCGTGTGCGCCGCAAGCTCTGCG 300
 DB 367 AGCAACCTGGCGCTGCTGGGCAACTCTGCGCGTGTGCGCCGCAAGCTCTGCG 426
 QY 301 CACCTGCGCGCAAGCGCGCGCGCAAGTGTCTCAAGACGCTATGAGCACTTCC 360
 DB 427 CACCTGCGCGCAAGCGCGCGCGCAAGTGTCTCAAGACGCTATGAGCACTTCC 486
 QY 361 AGCTCTCAAGAGAGAGAGAGGTAAGTGTGCGTGGCGCTGGGAGTCAAGATTTGCA 420
 DB 487 AGCTCTCAAGAGAGAGAGGTAAGTGTGCGTGGCGCTGGGAGTCAAGATTTGCA 546
 QY 421 AATTCTCAGCAGGCTAGTGGAGACAGGCTCCCCCAACCAAGCGGCAAGCGGAGT 480
 DB 547 AATTCTCAGCAGGCTAGTGGAGACAGGCTCCCCCAACCAAGCGGCAAGCGGAGT 606
 QY 481 CGGGGCGCGCCAGTGTGTGTGCGAGACGCGCGGAGAGAGGCGCCAGCGGCAAG 540
 DB 607 CGGGGCGCGCCAGTGTGTGTGCGAGACGCGCGGAGAGAGGCGCCAGCGGCAAG 666
 QY 541 CAGCAGTCAAGAGCCCGGCAAGCTTCTCTTGAAGCAAGTGAAGTGTGATCTCGGCTC 600
 DB 667 CAGCAGTCAAGAGCCCGGCAAGCTTCTCTTGAAGCAAGTGAAGTGTGATCTCGGCTC 726
 QY 601 CGGGTTGAGAGAGTACTGAGCATGAGGCGCCTTGGAGCAGGGGCTGAGTCCCG 660
 DB 727 CGGGTTGAGAGAGTACTGAGCATGAGGCGCCTTGGAGCAGGGGCTGAGTCCCG 786
 QY 661 CGGGCCAGAGGCTGTGGCGGCAAGTGTGTTGGGAGGCGCACGAGTGTCTGCGC 720

DB 787 CGGGCCAGAGGCTGTGGCGGCGGCAAGTGTGTTGGGAGGCGCACCGGAGTGTGCGC 846
 QY 721 TCAAGGAGCTTGGGCTGTGTGTTGATCATCAAGTTCTCAAGAGCTCTCTATCTGAG 780
 DB 847 TCAAGGAGCTTGGGCTGTGTGTTGATCATCAAGTTCTCAAGAGCTCTCTATCTGAG 906
 QY 781 GCCTTGTGGGCGCATCTAAGTGTGGCGCCCTGTGCAAGGCTCTGCGGGGCTGTCTTG 840
 DB 907 GCCTTGTGGGCGCATCTAAGTGTGGCGCCCTGTGCAAGGCTCTGCGGGGCTGTCTTG 966
 QY 841 ACTGAGGCGCTGAGAGAGGCTGTGTGGCGGAGGCTGTGCGCTGTGATGTGGAT 900
 DB 967 ACTGAGGCGCTGAGAGAGGCTGTGTGGCGGAGGCTGTGCGCTGTGATGTGGAT 1026
 QY 901 GAGGCTGAC 909
 DB 1027 GAGGCTGAC 1035

RESULT 5

AAD40080
 ID AAD40080 standard; DNA; 1924 BP.

XX AAD40080;
 XX 22-OCT-2002 (first entry)

DE Human DED4 (death effector domain) gene.

XX Human; death domain; DB; death effector domain; DED; Chlamydia infection;
 KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
 KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
 KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
 KW immunosuppressive; gene therapy; antisense therapy; gene; ds.

XX Homo sapiens.

OS Location/Qualifiers

XX Key 91..1044

FT CDS

FT /tag= a

FT /product= "Human DED4"

FT /note= "No stop codon"

FT /partial

FT misc_feature

FT 157..222

FT /tag= b

FT /note= "Nuclear localisation sequence"

XX WO200240680-A2.

XX 23-MAY-2002.

XX 15-NOV-2001; 2001NO-US044844.

XX 17-NOV-2000; 2000US-00715893.

XX 29-JUN-2001; 2001US-0301889P.

XX (BURN-) BURHAM INST.

XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Steiner-Hewen F;
 XX WPI: 2002-500222/53.
 DR P-PSDB; AAE24860.
 XX

XX New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.

PS Claim 19; Page 184-186; 209pp; English.

XX The invention relates to an isolated polypeptide comprising a death
 CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
 CC is useful for identifying a binding agent, preferably a protein or a drug

XX WPI; 2002-500222/53.
DR P-PSDB; AAE38903.
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
XX allergy, sepsis, autoimmunity, allograft rejection and other diseases.
PS Claim 19; Page 42-44; 99p; English.
XX
CC The present invention provides novel death domain (DD) and death effector
CC domain (DED) proteins and nucleic acids encoding them. The invention also
CC provides death domain containing protein such as Chlamydia trachomatis
CC death domain containing protein (CTDP) DD and neutral growth factor
CC receptor-interacting death domain (NIDD) DD. The invention is useful for
CC identifying a binding agent (e.g. protein or drug) that binds a DD, DED
CC or NB-ARC domain from DRP3, TRAF4, CTDP, DED4 or NIDD with a candidate
CC binding agent and identifying an effective agent (e.g. protein or drug)
CC that modulates the association of a DD, DED or NB-ARC domain with protein
CC that binds the DD, DED or NB-ARC domain. The invention is also useful for
CC modulating the level of cell process such as apoptosis, cell adhesion,
CC cell proliferation, cell stress responses, responses to microbial
CC infection and B cell immunoglobulin class switching. DDs, DEDs and NB-ARC
CC domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are
CC useful for discovery of drugs that suppress infection, autoimmunity,
CC inflammation, allergy, allograft rejection, sepsis and other diseases.
CC DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
CC autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
CC hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
CC inflammatory hyperplasia and smooth muscle cell proliferation in arteries
CC following balloon angioplasty (restenosis). The invention is also used in
CC antibody therapy and gene therapy. The present sequence is human DED4
CC full length gene. The DED4 gene is located on chromosome 19
XX
SQ Sequence 1924 BP; 360 A; 599 C; 594 G; 368 T; 0 U; 3 Other;
Query Match 94.4%; Score 858; DB 6; Length 1924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCGGGTGCACCCCGCGCTGCTGGAGAGATGAGTCTTGACTAC 60
DB 91 ATGCGCGTATCGGGTGCACCCCGCGCTGCTGGAGAGATGAGTCTTGACTAC 150
QY 61 TAGCGGATGCTGCTGCTTCAACGTATGTTGAGGTGCTGGCGGCACTACGAGTGC 120
DB 151 TAGCGGATGCTGCTGCTTCAACGTATGTTGAGGTGCTGGCGGCACTACGAGTGC 210
QY 121 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGCTCTGCGCGCGGAGGCTTAC 180
DB 211 GAGCTGAGCTCTGCTGCTTCTGCTGATGAGCTCTGCGCGCGGAGGCTTAC 270
QY 181 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGCACTGCGCGAG 240
DB 271 CGGGCCCGCAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCGCGGCACTGCGCGAG 330
QY 241 AGCAACCTGCGGCTGCTGGGCACTCTGCGCGCTGCTGCGCGGCACTGCTGCGG 300
DB 331 AGCAACCTGCGGCTGCTGGGCACTCTGCGCGCTGCTGCGCGGCACTGCTGCGG 390
QY 301 CACCTGCGCGCAGCGGCGCGGCACTGCTGCGAGAGCTATAGTATGAGCACTCC 360
DB 391 CACCTGCGCGCAGCGGCGCGGCACTGCTGCGAGAGCTATAGTATGAGCACTCC 450
QY 361 AGCTCTTCAAG 420
DB 451 AGCTCTTCAAG 510
QY 421 AATTCTCAGCAGGCTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 511 AATTCTCAGCAGGCTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
QY 481 CGGGGCGGCGCAGTGTGTGTGTCAGAGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 540

DB 571 CGGGGCGGCGCAGTGTGTGTGTCAGAGCGGCGGAGAGAGAGAGAGAGAGAGAG 630
QY 541 CAGCAGTCAAG 600
DB 631 CAGCAGTCAAG 690
QY 601 CGGGTTCAG 660
DB 691 CGGGTTCAG 750
QY 661 CGGGCCCGCAGGCGCTGCGCGGCGAGCTGAGAGTGTGTTGGAGAGAGAGAGAGAG 720
DB 751 CGGGCCCGCAGGCGCTGCGCGGCGAGCTGAGAGTGTGTTGGAGAGAGAGAGAGAG 810
QY 721 TCAAGGAG 780
DB 811 TCAAGGAG 870
QY 781 GCTTCTGGGGCGAGTACTGAGTGGCGGCGCTGCTGAGAGAGAGAGAGAGAGAGAG 840
DB 871 GCTTCTGGGGCGAGTACTGAGTGGCGGCGCTGCTGAGAGAGAGAGAGAGAGAGAG 930
QY 841 ACTGAGGCGCTGCGAG 900
DB 931 ACTGAGGCGCTGCGAG 990
QY 901 GAGGCTGAC 909
DB 991 GAGGCTGAC 999
RESULT 7
ID ADQ86891 standard; cDNA; 1979 BP.
AC ADQ86891;
DT 07-Oct-2004 (first entry)
XX Human tumour-associated antigenic target (TAT) cDNA sequence #3766.
DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
OS WO2004060270-A2.
PN 22-Jul-2004.
PD 15-Oct-2003; 2003WO-US029126.
PF 18-Oct-2002; 2002US-0418988P.
PR (GENT) GENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
PI Wu TD, Zhou Y;
XX WPI; 2004-534300/51.
DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
FT preventing or treating cell proliferative disorders such as cancer.
PS Claim 1; SEQ ID NO 3766; 5504bp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-

CC The invention relates to the isolation of genes encoding 9 human
CC apoptosis-related proteins. The nucleotide sequences AAG95790-A95798
CC encode the human apoptosis related proteins AAB15551-B15559. The genes
CC can be used to generate fusion proteins by linking to the gene for the
CC human immunoglobulin G Fc (IgG Fc) portion (AAG95799) for increasing the
CC stability of the fusion protein as compared to the human protein only.
CC The gene and encoded protein may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate apoptosis associated
CC protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,
CC melanomas and lymphomas), inflammation, autoimmune disorders (e.g.
CC multiple sclerosis) and viral infections (e.g. herpes)

SQ Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;

Query Match 94.4%; Score 858; DB 3; Length 2045;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY		ATGGGGCTATCGGGGGTCGACCCCGGCCCGGTGCTGGAGAGAGGATATGAGTCCCTGACCTAC	60
Db	121	ATGGGGCTATCGGGGGTCGACCCCGGCCCGGTGCTGGAGAGAGGATATGAGTCCCTGACCTAC	180
OY	61	TACGGGATGCTGTGCTTACCGTATGTTTGAAGTGTGGGCGGGGCACTGACCGAGTGC	120
Db	181	TACGGGATGCTGTGCTTACCGTATGTTTGAAGTGTGGGCGGGGCACTGACCGAGTGC	240
OY	121	GAGCTGGAAGCTCTCTGGCCCTTTCTGCTGATGAGGCTTCCTGGCGCCCGGAGGCTTAAGCC	180
Db	241	GAGCTGGAAGCTCTCTGGCCCTTTCTGCTGATGAGGCTTCCTGGCGCCCGGAGGCTTAAGCC	300
OY	181	CGGGGCCGAGCGGCTAGAGCTCCGTCAGCTGAGCTGAGCGCCCGCGGGGCAAGTGGCGGAG	240
Db	301	CGGGGCCGAGCGGCTAGAGCTCCGTCAGCTGAGCTGAGCGCCCGCGGGGCAAGTGGCGGAG	360
OY	241	AGCAACCTGCGGCTGTGGGGCACTCTTGCAGCTGCTGGCCCGCCGACGACCTGCTGCGG	300
Db	361	AGCAACCTGCGGCTGTGGGGCACTCTTGCAGCTGCTGGCCCGCCGACGACCTGCTGCGG	420
OY	301	CACCTGGCGCGCAAGCGGCGCGGCCCAAGTGTCTCCAGAACGCTATATGGTACCTTCC	360
Db	421	CACCTGGCGCGCAAGCGGCGCGGCCCAAGTGTCTCCAGAACGCTATATGGTACCTTCC	480
OY	361	AGCTCTTCAAAAGAGGACAGAGGGTAGCTCCGTCGCGCATCGGACATCAAGCACTTCTGCA	420
Db	481	AGCTCTTCAAAAGAGGACAGAGGGTAGCTCCGTCGCGCATCGGACATCAAGCACTTCTGCA	540
OY	421	AATTCTCAGCAGGGTCTAGTGGAGAGCAGGCTCCCGCCCAACCAAGCGGACCGGCGGAGT	480
Db	541	AATTCTCAGCAGGGTCTAGTGGAGAGCAGGCTCCCGCCCAACCAAGCGGACCGGCGGAGT	600
OY	481	CGGGGCCCGGCCCAAGTGTGTGTGTCANAGCGGCGCGGAGAGGGGGCCCCAGTCCGACCCCA	540
Db	601	CGGGGCCCGGCCCAAGTGTGTGTGTCANAGCGGCGCGGAGAGGGGGCCCCAGTCCGACCCCA	660
OY	541	CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAGGAAAGTGACTGTGACATCCGAGCTC	600
Db	661	CAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAGGAAAGTGACTGTGACATCCGAGCTC	720
OY	601	CGGGTTCAGAGAGTACTGCGAGCATGGGCGACGCTTGGAGGACGGGCGTGGCATCCCG	660
Db	721	CGGGTTCAGAGAGTACTGCGAGCATGGGCGACGCTTGGAGGACGGGCGTGGCATCCCG	780
OY	661	CGGGCCCAAGGCGCTGCGCGGCGAGCTGAGACGTGTTTGGGCAAGGCCACCGAGTGCACG	720
Db	781	CGGGCCCAAGGCGCTGCGCGGCGAGCTGAGACGTGTTTGGGCAAGGCCACCGAGTGCACG	840
OY	721	TCAAGGAGCCTTGGGCTGTGTGTTTGTGACATCAAAATTCTCAGAGCTCTCTATCTGGAC	780
Db	841	TCAAGGAGCCTTGGGCTGTGTGTTTGTGACATCAAAATTCTCAGAGCTCTCTATCTGGAC	900
OY	781	GCTTCTGTGGGCGACTACTTGAAGTGGCGCCCTGTGTCAGAGCCCTGTGGGGCGTGTCTTG	840
Db	901	GCTTCTGTGGGCGACTACTTGAAGTGGCGCCCTGTGTCAGAGCCCTGTGGGGCGTGTCTTG	960

QY	841	ACTAGAGCCCGCAGAGGCTGTGGCCCGGAGGCTGTCGCTGCTGCTGAGTGGAT	900
DB	961	ACTAGAGCCCTCCAGAGGCTGTGGCCCGGAGGCTGTTCCGCTCTGGTCAGTGTGGAT	1020
QY	901	GAGGCTGAC	909
DB	1021	GAGGCTGAC	1029

RESULT 9

ID AAH33283 standard; cDNA; 2045 BP.

AC AAH33283;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen encoding cDNA SEQ ID NO:339.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX

XX
XX

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

PR 03-NOV-1999; 99US-0163280P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse

DR WPI: 2001-235357/24.
DR P-PSDB: AAC73953

XX
DT

PT useful for prevention

XX
PS
XX
Claim 1; Page 2451; 9803pp; English.

CC cancer-associated nucleic acid molecules (N) and proteins (P) where t
CC AAH32943 to AAH37195 and AAG73514 to AAG7788 represent human colon

CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy

CC and vaccine production. N and P may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate P expression.

expression by rectifying mutations or deletions in a patient's genome

CC supplement the patients own production of P. Additionally, N may be us
CC that direct the activity of P by expressing inactive proteins or co

CC into a host cell and culturing the cell to express the proteins. N and CC to produce the colon cancer-associated F8, by inserting the nuclear ac

CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent

CC sequences used in the exemplification of the present invention. N.B.:
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at

CC to 1052, 7921 and 7922

Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;

Query Match 94.4%; Score 858; DB 4; Length 2045;

Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps

QY 1 ATGGCGCTATCCGGGTCGACCCCGGCCCTGCTGGAGGAGTGAAGTGCCTGACTAC

```

Db      121 ATGGCCCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGATGATGCTTGAATAC 180
Qy      61 TACGGATGCTGTCTGCTTCAACCGTATGTTGAGAGTGGTGGCGGCAATGACCGAGTGC 120
Db      181 TACGGATGCTGTCTGCTTCAACCGTATGTTGAGAGTGGTGGCGGCAATGACCGAGTGC 240
Qy      121 GAGCTGGAGTCTCTGGCTTTTCTGCTGATGAGGCTTCCTGGCCCGCCGAGGCTTAAGCC 180
Db      241 GAGCTGGAGTCTCTGGCTTTTCTGCTGATGAGGCTTCCTGGCCCGCCGAGGCTTAAGCC 300
Qy      181 CGGGCCCGAGCGGCTAGAGCTCTGCTGAGCTGGAGCGCGCGGAGAGTGGCGAG 240
Db      301 CGGGCCCGAGCGGCTAGAGCTCTGCTGATGAGGCTTCCTGGCCCGCCGAGGAG 360
Qy      241 AGCAACTGGGCTGCTGGGCAACTCTGCGGCTGGCTGGCCCGGCAAGACTGCTGC 300
Db      361 AGCAACTGGGCTGCTGGGCAACTCTGCGGCTGGCTGGCCCGGCAAGACTGCTGC 420
Qy      301 CACCTGGCGGCAAGCGGCGCGGCAAGTCTTCAGAGCTATGCTATGGCACTCC 360
Db      421 CACCTGGCGGCAAGCGGCGCGGCAAGTCTTCAGAGCTATGCTATGGCACTCC 480
Qy      361 AGCTCTTCAAGAGCAGAGGGTATGCTGCGGCTGGCTGGCCCGGCAAGCTATGCA 420
Db      481 AGCTCTTCAAGAGCAGAGGGTATGCTGCGGCTGGCTGGCCCGGCAAGCTATGCA 540
Qy      421 AATTTCTCAGCAGGGTCAGTGGAGAGCAGGCTCCCGCCCAACAGCGGCGGAGT 480
Db      541 AATTTCTCAGCAGGGTCAGTGGAGAGCAGGCTCCCGCCCAACAGCGGCGGAGT 600
Qy      481 CGGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db      601 CGGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy      541 CACGAGTCAGAGCGGCGGCAAGCTTCTGGAAGCAAAAGTGAAGCTGTGATCCGCTC 600
Db      661 CACGAGTCAGAGCGGCGGCAAGCTTCTGGAAGCAAAAGTGAAGCTGTGATCCGCTC 720
Qy      601 CGGGTTCGAGCAGAGTACTGCGAGCATGAGGCGACCTTGGAGCAGGGCGTGGCATCC 660
Db      721 CGGGTTCGAGCAGAGTACTGCGAGCATGAGGCGACCTTGGAGCAGGGCGTGGCATCC 780
Qy      661 CGGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db      781 CGGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy      721 TCAAGGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db      841 TCAAGGAGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy      781 GCCTTCTGGGCGCACTAAGTGGGCGCTGCTGAGGCGCTGCGGGCGTGTTCCTG 840
Db      901 GCCTTCTGGGCGCACTAAGTGGGCGCTGCTGAGGCGCTGCGGGCGTGTTCCTG 960
Qy      841 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGAGGCTTGGCTGCTGCTGCTGCTG 900
Db      961 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGAGGCTTGGCTGCTGCTGCTGCTG 1020
Qy      901 GAGGCTGAC 909
Db      1021 GAGGCTGAC 1029

```

RESULT 10
AAFI8296 standard; DNA, 2044 BP.

AAFI8296;

14-MAR-2001 (first entry)

Lung cancer associated polynucleotide sequence SEQ ID 315.

```

KW      Human; lung cancer associated protein; neuroprotective; cytosolic;
KW      cardioactive; immunomodulatory; muscular active; vulnery;
KW      gastrointestinal; nephrotoxic; antineoplastic; gynecological;
KW      antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KW      proliferative disorder; wound healing; infectious disease; ds.
XX
OS      Homo sapiens.
PN      WO20005180-A2.
PD      21-SEP-2000.
PF      08-MAR-2000; 2000MO-US005918.
PR      12-MAR-1999; 99US-0124270P.
PA      (HUMA-) HUMAN GENOME SCI INC.
PA      (ROSE/) ROSEN C A.
XX
PI      Ruben SM;
XX
DR      WPI; 2000-587514/55.
DR      P-PDB; AAB58420.
XX
PT      Lung cancer associated gene sequences, referred to as lung cancer
PT      antigens, useful for treatment, prevention, and diagnosis of disorders
PT      such as lung cancer.
PS      Claim 1; Page 772-773; 1425pp; English.
XX
CC      Polynucleotide sequences AAF18424 - AAF18424 encode human lung cancer
CC      associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC      associated proteins and polynucleotide sequences, their agonists, and
CC      antagonists may have neuroprotective; cytostatic; cardioactive;
CC      immunomodulatory; muscular active general; vulnery; gastrointestinal
CC      general; nephrotoxic; antineoplastic; gynecological; or antibacterial
CC      activity. The invention also includes antibodies specific for the protein
CC      or polynucleotide sequences. The lung cancer associated polynucleotide
CC      sequences may be used for detection of lung cancer, chromosome
CC      identification, as chromosome markers, and for numerous other diagnostic
CC      or research purposes. The proteins may be used to treat disorders such as
CC      neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC      cardiovascular, renal, and proliferative disorders. The proteins may also
CC      be used in the treatment of wounds and infectious diseases.
CC      Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC      used in the course of the invention for the identification and
CC      characterization of the polynucleotide and protein sequences
XX
SQ      Sequence 2044 BP; 457 A; 620 C; 603 G; 364 T; 0 U; 0 Other;

```

Query Match 84.5%; Score 768; DB 3; Length 2044;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 818; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 ATGGCCCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGATGATGCTTGAATAC 60
Db      121 ATGGCCCTATCCGGGTCGACCCCGCCCGCTGCTGGAGAGATGATGCTTGAATAC 180
Qy      61 TACGGATGCTGTCTGCTTCAACCGTATGTTGAGAGTGGTGGCGGCAATGACCGAGTGC 120
Db      181 TACGGATGCTGTCTGCTTCAACCGTATGTTGAGAGTGGTGGCGGCAATGACCGAGTGC 240
Qy      121 GAGCTGGAGTCTCTGGCTTTTCTGCTGATGAGGCTTCCTGGCCCGCCGAGGCTTAAGCC 180
Db      241 GAGCTGGAGTCTCTGGCTTTTCTGCTGATGAGGCTTCCTGGCCCGCCGAGGCTTAAGCC 300
Qy      181 CGGGCCCGAGCGGCTAGAGCTCTGCTGAGCTGGAGCGCGGAGAGTGGCGAG 240
Db      301 CGGGCCCGAGCGGCTAGAGCTCTGCTGATGAGGCTTCCTGGCCCGCCGAGGAG 360
Qy      241 AGCAACTGGGCTGCTGGGCAACTCTGCGGCTGGCTGGCCCGGCAAGACTGCTGC 300
Db      361 AGCAACTGGGCTGCTGGGCAACTCTGCGGCTGGCTGGCCCGGCAAGACTGCTGC 420

```

QY 301 CACCTGGCGCGCAGAGGGCGCCCGGCCAAGTGTCTCCGAAAGCTATATGCTATGGCACTCC 360
 Db 421 CACCTGGCGCGCAGAGGGCGCCCGGCCAAGTGTCTCCGAAAGCTATATGCTATGGCACTCC 480
 QY 361 AGCTCTTCAAAGGAGCAGAGGGTACTGCGCGCGATCGGAGATCAAGCAATTCGCA 420
 Db 481 AGCTCTTCAAAGGAGCAGAGGGTACTGCGCGCGATCGGAGATCAAGCAATTCGCA 540
 QY 421 AATTCTCAGCAGAGGGTACTGCGGAGAGAGGCTCCCGCCCAACAAAGCGGACGCGCGAAGT 480
 Db 541 AATTCTCAGCAGAGGGTACTGCGGAGAGAGGCTCCCGCCCAACAAAGCGGACGCGCGAAGT 600
 QY 481 CGGGGCGCGGCCCAAGTGTGTGTGCAGACGAGCGGCGGAGAGAGGGGCCCAAGCCGACCCGAC 540
 Db 601 CGGGGCGCGGCCCAAGTGTGTGTGCAGACGAGCGGCGGAGAGAGGGGCCCAAGCCGACCCGAC 660
 QY 541 CAGCAGTCAAGAGCCCGCCAGACTTCTCTGAAAGGCAAAAGTGAACCTGTGCATCCGACTC 600
 Db 661 CAGCAGTCAAGAGCCCGCCAGACTTCTCTGAAAGGCAAAAGTGAACCTGTGCATCCGACTC 720
 QY 601 CGGGTTCGAGCAGAGTACTGCGAGCATGGGCGACGCTTGAGAGCAGGGGCGTGGATCCCGG 660
 Db 721 CGGGTTCGAGCAGAGTACTGCGAGCATGGGCGACGCTTGAGAGCAGGGGCGTGGATCCCGG 780
 QY 661 CGGCCCCAGGCCCTGGCGCGGCAGCTGAGCGATGTTTGGGCGAGGCCCAACCGCAGTGTGCGC 720
 Db 781 CGGCCCCAGGCCCTGGCGCGGCAGCTGAGCGATGTTTGGGCGAGGCCCAACCGCAGTGTGCGC 840
 QY 721 TCAAGGGACCTGGGCTCTGTGTGTTTGTGACATCAAGATTCTCAGAGCTCTTCTATCTGAC 780
 Db 841 TCAAGGGACCTGGGCTCTGTGTGTTTGTGACATCAAGATTCTCAGAGCTCTTCTATCTGAC 900
 QY 781 GCGTTCTGGGGCGACTACTGAGTGGCGCCCTGCTCGAG 819
 Db 901 GCGTTCTGGGGCGACTACTGAGTGGCGCCCTGCTCGAG 939

```

RESULT 11
AAS62603
ID AAS62603 standard; cDNA; 1966 BP.
XX
XX
AC AAS62603;
XX
XX
DT 14-FEB-2002 (first entry)
XX
DE cDNA sequence #390 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antirheumatic; ss.
XX
OS Homo sapiens.
XX
PN WO200177291-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010485.
XX
PR 06-APR-2000; 2000US-0195604P.
XX
PA (GENY ) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulakota K, Graham JR;
XX
DR WPI; 2002-010900/01.
PT New polynucleotides encoding secreted proteins useful for treating e.g.
XX asthma, HIV and Crohn's disease.
XX

```

PS Claim 1, Page 280; 391pp; English.

XX

CC The present invention relates to the isolation of novel cDNA sequences

CC which encode human secreted proteins. The cDNA sequences have been

CC derived from a variety of human tissues. The invention also provides a

CC method for producing proteins from these polynucleotide sequences. The

CC proteins are useful for identifying compounds that modulate their

CC activity and production, and the cell is also useful for identifying

CC compounds that modulate expression of the polynucleotide sequences

CC encoding the secreted proteins. The sequences of the invention are useful

CC for treating diseases such as hyperproliferative disorders (e.g. cancer),

CC immune deficiency disorders (e.g. severe combined immunodeficiency

CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders

CC (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis) and

CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of

CC the invention are also useful in gene therapy. AAS62214-AAS62838

CC represent the cDNA sequences of the invention that encode for novel human

CC secreted proteins

XX

SQ Sequence 1966 BP; 356 A; 618 C; 614 G; 378 T; 0 U; 0 Other;

Query Match	59.2%	Score 538,	DB 6;	length 1966;
Best Local Similarity	99.8%	Pred. No. 2.9e-236;		
Matches 588; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	ATGAGGCTATCCGGGTGCACCCCGGCCCCCGTCTGGAGAGAGATGACGTCCCTGACCTAC	60
Db	50	ATGGGGCTATCCGGGTGCACCCCGGCCCCCGTCTGGAGAGAGATGACGTCCCTGACCTAC	109
QY	61	TACGGGATGCTGTCCGCTTCAACCGTATGTTTCGAGTGTGATGGCGGGCAACTGACCGAGTGC	120
Db	110	TACGGGATGCTGTCCGCTTCAACCGTATGTTTCGAGTGTGATGGCGGGCAACTGACCGAGTGC	169
QY	121	GAGCTGAGAGCTCTGAGCCCTTCTGCTGATGAGGCTCTGAGCGCGCCGAGAGCTTTAGCC	180
Db	170	GAGCTGAGAGCTCTGAGCCCTTCTGCTGATGAGGCTCTGAGCGCGCCGAGAGCTTTAGCC	229
QY	181	CGGGGCCCCGAGGGGCTAGAGCTCTGCTGAGGCTGAGAGCGCGGGGGGAGATGGCGGAG	240
Db	230	CGGGGCCCCGAGGGGCTAGAGCTCTGCTGAGGCTGAGAGCGCGGGGGGAGATGGCGGAG	289
QY	241	AGCAACTGCGGGCTCTGGGGGCAACTCTGACCGTGTGGCCCGGCAACGACTGTGCGCG	300
Db	290	AGCAACTGCGGGCTCTGGGGGCAACTCTGACCGTGTGGCCCGGCAACGACTGTGCGCG	349
QY	301	CACCTGGCGCGCAAGCGGCGCGGCCAGTGTCTCGAAGACGCTATAGCTATGCGACCTCC	360
Db	350	CACCTGGCGCGCAAGCGGCGCGGCCAGTGTCTCGAAGACGCTATAGCTATGCGACCTCC	409
QY	361	AGCTCTTCAMAAAGACAGAGGGTATGCTGCCGTCCGCTGGCAGTCAAGCATTTTGA	420
Db	410	AGCTCTTCAMAAAGACAGAGGGTATGCTGCCGTCCGCTGGCAGTCAAGCATTTTGA	469
QY	421	AATTCACAGCAGGTCAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGAGCGCGAGT	480
Db	470	AATTCACAGCAGGTCAGTGGGAGACAGGCTCCCCCAACCAAGCGGCGAGCGCGAGT	529
QY	481	CGGGGCGCGGCCAGTGTGTGTCCAGACGGCGCGAGAGAGGGGCCCCAGCCGCAACCCAG	540
Db	530	CGGGGCGCGGCCAGTGTGTGTCCAGACGGCGCGAGAGAGGGGCCCCAGCCGCAACCCAG	589
QY	541	CAGCATCAGAGCCCGCCAGACCTTCTCTTAAGCAAAAGTACCTGTG	589
Db	590	CAGCATCAGAGCCCGCCAGACCTTCTCTTAAGCAAAAGTACCTGTG	638

RESULT 12	
AAH99646	
ID	AAH99646 standard; cDNA; 1570 BP.
XX	
AC	
XX	AAH99646;
XX	
DT	16-OCT-2001 (first entry)

XX Human protein encoding cDNA sequence SEQ ID NO:481.
XX
XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KM antibacterial; endocrine; cardiac; central nervous system; virucide;
KM anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anemica;
KM antiaggregant; haemostatic; vulnery; antileuc; osteopathic; eczema;
KM dermatologic; antiallergic; antiaesthetic; antidiabetic; cytostatic;
KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KM anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KM allergic rhinitis; diabetes; multiple sclerosis; depression;
KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KM neurological disorder; ss.
XX
XX Homo sapiens.
OS
PN W0200153455-A2.
XX
XX 26-JUL-2001.
PD
XX 22-DEC-2000; 2000WO-US035017.
PF
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-0048725.
PR
XX 25-APR-2000; 2000US-00552317.
PR
XX (HXSE-) HXSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI: 2001-457603/49.
DR
XX P-PSDB: AAM25705.
PT
XX Isolated human polynucleotides encoding polypeptides, useful for the
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
XX Claim 1; Page 552; 1217pp; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnery;
CC antileuc; osteopathic; dermatologic; antiallergic; antiaesthetic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
XX Sequence 1570 BP; 311 A; 488 C; 466 G; 305 T; 0 U; 0 Other;
SQ

Query Match 58.4%; Score 531; DB 4; Length 1570;
Best Local Similarity 99.8%; Pred. No. 4.7e-233;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY GTGCTCCAGAACGCTATAGTACCTCCAGCTCTTCAAGAGAGACAGAGGTAGC 387
|||||

Db 76 GTGCTCCAGAACGCTATAGTACCTCCAGCTCTTCAAGAGAGACAGAGGTAGC 135
QY 388 TGCCGTGCGCGTCGCGAGTCAGACAGTTCTGCAAAATTCAGACAGGCTAGTGGAGACA 447
Db 136 TGCCGTGCGCGTCGCGAGTCAGACAGTTCTGCAAAATTCAGACAGGCTAGTGGAGACA 195
QY 448 GGCCTCCCCCAACCAAGCGGCGAGCGCGGAGTCGGGGCCGGCCAGTGGTGGCCAGA 507
Db 196 GGCCTCCCCCAACCAAGCGGCGAGCGCGGAGTCGGGGCCGGCCAGTGGTGGCCAAA 255
QY 508 CGCGCGCGAGAGAGGCGGCCCGCAGCGCACCCACAGACAGTCAGAGCCCGCAGACTTCC 567
Db 256 CGCGCGCGAGAGAGGCGGCCCGCAGCGCACCCACAGACAGTCAGAGCCCGCAGACTTCC 315
QY 568 TCTGAAGGCAAAAGTACCTGTGACATCGGCTCCGGGTTTGAGCAGATCTGCGAGCAT 627
Db 316 TCTGAAGGCAAAAGTACCTGTGACATCGGCTCCGGGTTTGAGCAGATCTGCGAGCAT 375
QY 628 GGGCGACCTTTGAGCAGCGGGCGTGGCATTCGGCGGCCGCCAGGCGCTGGCGCGAGCTG 687
Db 376 GGGCGACCTTTGAGCAGCGGGCGTGGCATTCGGCGGCCGCCAGGCGCTGGCGCGAGCTG 435
QY 688 GACGTGTTTGGGAGAGGCGCACCGCAGTGTGCTGCTCAAGGAGCTGGGCTCTGTGTTGT 747
Db 436 GACGTGTTTGGGAGAGGCGCACCGCAGTGTGCTGCTCAAGGAGCTGGGCTCTGTGTTGT 495
QY 748 GACATCAAGTTCTCAGAGCTCTCTATCTTGAGCGCTTCTGGGGCGACTGAGTGGC 807
Db 496 GACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGGCGACTGAGTGGC 555
QY 808 GCCCTGTCGAGGCGCCCTGGCGGGCGGTTCCTGATCGAGGCGCTCGAGAGGCTGTGGGC 867
Db 556 GCCCTGTCGAGGCGCCCTGGCGGGCGGTTCCTGATCGAGGCGCTCGAGAGGCTGTGGGC 615
QY 868 CGGAGGCTGTGCGCTGCTGTGTCAGTGTGATGAGGCTGAC 909
Db 616 CGGAGGCTGTGCGCTGCTGTGTCAGTGTGATGAGGCTGAC 657

RESULT 13
AAH07925
ID AAH07925 standard; cDNA; 626 BP.
XX
XX AAH07925;
AC
XX 26-JUN-2001 (first entry)
DT
XX Human cDNA clone (5'-primer) SEQ ID NO:4760.
DE
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
KM
XX
OS
XX Homo sapiens.
XX
PN EPI074617-A2.
PD
XX 07-FEB-2001.
PF
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PA
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saio K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
DR
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT

QY 181 CGGAGCCGCGAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGGCGAGTGGCGGAG 240
 CC |||||
 CC CGGAGCCGCGAGCGGCTTGAAGCTCTGCTGAGCTGAGCGCGGCGAGTGGCGGAG 716
 CC |||||
 QY 241 AGCAACCTGGGCTGCTGGGGGCAACTCTGCGGTGCTGGCCCGGCAAGCACTGGCTGG 300
 CC |||||
 Db 715 AGCAACCTGGGCTGCTGGGGGCAACTCTGCGGTGCTGGCCCGGCAAGCACTGGCTGG 656
 CC |||||
 QY 301 CACCTGGCGCGCAAGCGGCGCGCGCAAGTCTTCAGAGAGCTATAGCTATGGCACTTCC 360
 CC |||||
 Db 655 CACCTGGCGCGCAAGCGGCGCGCGCAAGTCTTCAGAGAGCTATAGCTATGGCACTTCC 596
 CC |||||
 QY 361 AGCTCTTCAAGAGAGAGAGGAGTGTGCTGCGGTGCGGCTGGCAAGTCTTGGCA 420
 CC |||||
 Db 595 AGCTCTTCAAGAGAGAGAGGAGTGTGCTGCGGTGCGGCTGGCAAGTCTTGGCA 536
 CC |||||
 QY 421 AATTCTCAGCAGG 434
 CC |||||
 Db 535 AATTCTCAGCAGG 522
 CC |||||

RESULT 15

AA158961 standard; cDNA; 1067 BP.

AA158961;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 1164.

Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukemia; ss.

Homo sapiens.

MO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US034263.

23-DEC-1999; 9US-00471275.

21-JAN-2000; 2000US-00488725.

25-APR-2000; 2000US-0052317.

20-JUN-2000; 2000US-00598042.

19-JUL-2000; 2000US-00620312.

03-AUG-2000; 2000US-00653450.

14-SEP-2000; 2000US-00662191.

19-OCT-2000; 2000US-00693036.

29-NOV-2000; 2000US-00727344.

(HYSE-) HYSEQ INC.

WPI; 2001-442253/47.

P-PSDB; AAM39805.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

Claim 1; SEQ ID NO 1164; 10078bp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM38642-AA42213) with neurotrophic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and CC localised neuropathies and central nervous system diseases, such as CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the CC utilisation of the activities such as: immune system suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders. Note: The sequence data for this patent did not form CC part of the printed specification

SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;

Query Match 30.9%; Score 281; DB 4; Length 1067;

Best Local Similarity 99.3%; Pred. No. 1.7e-118;

Matches 431; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGGCGCTATCCGGGTGACACCCGCGCCGCTGCGAGAGAGATAGTCTGCACTAC 60
 Db 151 ATGGCGCTATCCGGGTGACACCCGCGCCGCTGCGAGAGAGATAGTCTGCACTAC 210
 CC |||||
 QY 61 TACGGAGTGTGCTGCTTCAACCTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 120
 CC |||||
 Db 211 TACGGAGTGTGCTGCTTCAACCTATGTTGAGAGTGTGGCGGCACTGACCGAGTGC 270
 CC |||||
 QY 121 GAGCTGAGCTCTGCGCTTTCTGTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCC 180
 CC |||||
 Db 271 GAGCTGAGCTCTGCGCTTTCTGTGATGAGAGCTCTGAGCGCGCGGAGGCTTAGCC 330
 CC |||||
 QY 181 CGGAGCCGCGAGCGGCTTGAAGCTCTGAGCTGAGCGCGCGGCGAGTGGCGAG 240
 CC |||||
 Db 331 CGGAGCCGCGAGCGGCTTGAAGCTCTGAGCTGAGCGCGCGGCGAGTGGCGAG 390
 CC |||||
 QY 241 AGCAACCTGGGCTGCTGGGGCAACTCTGCGGTCTGCGCCCGGCAAGCACTGGCTGG 300
 CC |||||
 Db 391 AGCAACCTGGGCTGCTGGGGCAACTCTGCGGTCTGCGCCCGGCAAGCACTGGCTGG 450
 CC |||||
 QY 301 CACCTGGCGCGCAAGCGCGCGCGCAAGTCTTCAGAGAGCTATAGCTATGGCACTTCC 360
 CC |||||
 Db 451 CACCTGGCGCGCAAGCGCGCGCGCAAGTCTTCAGAGAGCTATAGCTATGGCACTTCC 510
 CC |||||
 QY 361 AGCTCTTCAAGAGAGAGAGGAGTGTGCTGCGGTGCGGAGTCAAGCACTTGGCA 420
 CC |||||
 Db 511 AGCTCTTCAAGAGAGAGAGGAGTGTGCTGCGGTGCGGAGTCAAGCACTTGGCA 570
 CC |||||
 QY 421 AATTCTCAGCAGG 434
 CC |||||
 Db 571 AATTCTCAGCAGG 584
 CC |||||

RESULT 16

ADQ99183 standard; cDNA; 1067 BP.

ADQ99183;

23-SEP-2004 (first entry)

DNA encoding human GPCR-like protein seqid 853.

ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;

KM antidiabetic; GPCR-like protein; ophthalmic disorder;

KM neurological disorder; immunological disorder; nephritic disorder;

KM hormonal dysfunction; cancer; atherosclerosis; diabetes;

KM molecular weight marker; food supplement; human; ss.

Homo sapiens.

US6569662-B1.

PD 27-MAY-2003.
XX 19-JUL-2000; 2000US-00620312.
XX 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Zhou P, Drmanac RT;
XX WPI; 2001-442255/47.
XX
XX New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT useful for treating diseases of ophthalmic, neurological, immunological
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT and diabetes.
XX
XX Example 2; SEQ ID NO 853; 92pp; English.
XX
XX The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (S1) of 749, 3188, 2484, 1169, 2936, 1467, 5774, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translated or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
XX polynucleotide of the invention.
XX
XX Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;

Query Match 30.9%; Score 281; DB 5; Length 1067;
Best Local Similarity 99.3%; Pred. No. 1.7e-118;
Matches 431; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCGGATATCGGGGTGACCCCGCCGCTGGGAGAGATGCTGACTAC 60
DB 151 ATGGCGGATATCGGGGTGACCCCGCCGCTGGGAGAGATGCTGACTAC 210
QY 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGAGCGGGCACTGACGAGTGC 120
DB 211 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGAGCGGGCACTGACGAGTGC 270
QY 121 GAGCTGAGTCTGCTGCTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAC 180
DB 271 GAGCTGAGTCTGCTGCTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAC 330
QY 181 CGGGCCCGGACCGGCTTGAAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 240
DB 331 CGGGCCCGGACCGGCTTGAAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 390
QY 241 AGCAACTGCGGGTGTGCTGAGAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 300
DB 391 AGCAACTGCGGGTGTGCTGAGAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 450
QY 301 CACCTGCGCGCAAGCGGCGCGGCACTGCTCAAGACGCTATAGCTATGCACTTCC 360
DB 451 CACCTGCGCGCAAGCGGCGCGGCACTGCTCAAGACGCTATAGCTATGCACTTCC 510
QY 361 AGCTCTTCAAGAGAGAGAGAGTGTGCTGCGGCTGCGGCACTCAAGAGTTCGCA 420
DB 511 AGCTCTTCAAGAGAGAGAGAGTGTGCTGCGGCTGCGGCACTCAAGAGTTCGCA 570
QY 421 AATTCTGACAGG 434
DB 571 AATTCTGACAGG 584

RESULT 17

ADB48943
ID ADB48943 standard; cDNA; 1067 BP.
XX
XX ADB48943;
AC
XX 04-DEC-2003 (first entry)
DT
XX Novel human cDNA SEQ ID NO 853.
DE
XX ss; cancer; neurodegenerative disease; human.
XX
XX Homo sapiens.
OS
XX US2003104529-A1.
PN
XX 05-JUN-2003.
PD
XX
XX 04-JAN-2002; 2002US-00037270.
PF
XX
XX 21-JAN-2000; 2000US-00488725.
PR
XX 25-APR-2000; 2000US-00552317.
PR 19-JUL-2000; 2000US-00620312.
XX
XX (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIU/) LIU C.
PA (ASUNDI/) ASUNDI V.
PA (DRMANAC/) DRMANAC R T.
XX
XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
PI WPI; 2003-678194/64.
XX
XX New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
XX
XX Claim 1; SEQ ID NO 853; 92pp; English.
XX
XX The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?docID=20030104529.
XX
XX Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;

Query Match 30.9%; Score 281; DB 9; Length 1067;
Best Local Similarity 99.3%; Pred. No. 1.7e-118;
Matches 431; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGGCGGATATCGGGGTGACCCCGCCGCTGGGAGAGATGAGTGTGCTGACTAC 60
DB 151 ATGGCGGATATCGGGGTGACCCCGCCGCTGGGAGAGATGAGTGTGCTGACTAC 210
QY 61 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGAGCGGGCACTGACGAGTGC 120
DB 211 TACGGGATGCTGTGCTTACCGTATGTTGAGAGTGAGCGGGCACTGACGAGTGC 270
QY 121 GAGCTGAGTCTGCTGCTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAC 180
DB 271 GAGCTGAGTCTGCTGCTTTCTGCTGATGAGGCTCTGCGCGCGGAGGCTTAC 330
QY 181 CGGGCCCGGACCGGCTTGAAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 240
DB 331 CGGGCCCGGACCGGCTTGAAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 390
QY 241 AGCAACTGCGGGTGTGCTGAGAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 300
DB 391 AGCAACTGCGGGTGTGCTGAGAGTCTGCTGAGAGTGAGCGCGGCACTGAGCGAG 450

```

Qy 301 CACCTGGCGGCGCAGCGCGCGCCAGTGTCTCCAGAACGCTATAGCTATGGACCTCC 360
Db 451 CACCTGGCGGCGCAGCGCGCGCCAGTGTCTCCAGAACGCTATAGCTATGGACCTCC 510
Qy 361 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 511 AGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
Qy 421 AATTCTCAGCAGGG 434
Db 571 AATTCTCAGCAGGG 584

RESULT 18
ADQ1808/c
XX ID ADQ1808 standard; DNA; 111084 BP.
XX AC ADQ1808;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 1627.
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnick A;
XX DR WPI; 2004-441208/41.
XX PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX PS Example 2; SEQ ID NO 1627; 210pp; English.
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX DNA of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 111084 BP; 24747 A; 31184 C; 30978 G; 24175 T; 0 U; 0 Other;

Query Match 30.5%; Score 277; DB 12; Length 111084;
Best Local Similarity 99.7%; Pred. No. 8.9e-117;
Matches 327; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGCGCATCCGGGTGACCCCGCCCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 8862 ATGGCGCATCCGGGTGACCCCGCCCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAG 8803

```

```

Qy 61 TAGGGATGCTGTGCTTCACTGATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 120
Db 8802 TAGGGATGCTGTGCTTCACTGATGTTGAGAGTGTGGCGGGCAACTGACCGAGTGC 8743
Qy 121 GAGCTGAGCTCTTGACCTTTCTGTGATGAGAGCTCTTGCGCGCCGAGAGCTTAGCC 180
Db 8742 GAGCTGAGCTCTTGACCTTTCTGTGATGAGAGCTCTTGCGCGCCGAGAGCTTAGCC 8683
Qy 181 CGGCGCCGCGAGCGGCTTAGAGCTCTTGAGAGCTGAGAGCGCGCGGCACTGCGCGAG 240
Db 8682 CGGCGCCGCGAGCGGCTTAGAGCTCTTGAGAGCTGAGAGCGCGCGGCACTGCGAG 8623
Qy 241 AGCAACTGCGGCTGCTGGGGCAACTCCGCGGCTGCTGGCGCGCACTGCTGCGCG 300
Db 8622 AGCAACTGCGGCTGCTGGGGCAACTCCGCGGCTGCTGGCGCGCACTGCTGCGCG 8563
Qy 301 CACCTGGCGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 328
Db 8562 CACCTGGCGGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 8535

RESULT 19
AAD40075
XX ID AAD40075 standard; cDNA; 303 BP.
XX AC AAD40075;
XX DT 22-OCT-2002 (first entry)
XX DE Human DED4 DED (death effector domain) cDNA.
XX KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
XX NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
XX inflammation; allergy; autoimmunity; allograft rejection; cell division;
XX immune-based pathology; fibrosis; arthritis; graft versus host disease;
XX immunosuppressive; gene therapy; antisense therapy; gene; ss.
XX OS Homo sapiens.
XX PN WO200240680-A2.
XX PD 23-MAY-2002.
XX PF 15-NOV-2001; 2001WO-US044844.
XX PR 17-NOV-2000; 2000US-00715893.
XX PR 29-JUN-2001; 2001US-0301889P.
XX PA (BURN-) BURHAM INST.
XX PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
XX PI Steiner-Biemer F;
XX DR WPI; 2002-500222/53.
XX DR P-PSDB; AAE24855.
XX PT New polypeptide comprising a death domain or death effector domain,
XX useful for discovery of drugs that suppress infection, inflammation,
XX PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX PS Claim 18; Page 174-175; 209pp; English.
XX CC The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC

```

CC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
 CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
 CC detecting the association of the domain and the candidate binding agent,
 CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
 CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
 CC spectroscopy (MS) and FPA. The invention is useful for modulating the
 CC level of a cell process such as cell proliferation, cell adhesion, cell
 CC stress responses, responses to microbial infection, and B cell
 CC immunoglobulin class switching, in particular apoptosis within a cell.
 CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
 CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
 CC CTDD DD protein is useful for detecting a Chlamydia infection. The
 CC invention is useful for modulating the activity of oncogenic proteins,
 CC for treating a pathology caused by the oncogenic proteins and for
 CC treating bacterial infections by modulating the activity of bacterial
 CC proteins. The protein and antibody specific for it are useful for
 CC discovery of drugs that suppress infection, inflammation, allergy,
 CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
 CC is useful for treating immune-based pathologies, pathologies associated
 CC with cell division, inflammatory diseases such as sepsis, fibrosis,
 CC arthritis, graft versus host disease. The invention is used in antisense
 CC therapy and gene therapy. The present sequence is human DED4 DED CDNA
 XX

SO Sequence 303 BP; 41 A; 96 C; 115 G; 51 T; 0 U; 0 Other;

Query Match 27.7%; Score 252; DB 6; Length 303;

Best Local Similarity 99.7%; Pred. No. 3.6e-105; Mismatches 1; Indels 0; Gaps 0;

Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 34 TGGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGTGCTTACCGTATGTTGAG 93

DB 1 TGGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGTGCTTACCGTATGTTGAG 60

QY 94 GTGGTGGGCGGCGCACTGACGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 153

DB 61 GTGGTGGGCGGCGCACTGACGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 120

QY 154 GCTTCTGCGCGCGCGGAGCTTACCCGCGCGCGGAGCTTACCGCTGCTGCTGAG 213

DB 121 GCTTCTGCGCGCGCGGAGCTTACCCGCGCGCGGAGCTTACCGCTGCTGCTGAG 180

QY 214 CTGAGAGCGCGCGCGGAGCTGCGGAGAGCAACTGCGGCTGTGCGGCAACTCTGCGC 273

DB 181 CTGAGAGCGCGCGCGGAGCTGCGGAGAGCAACTGCGGCTGTGCGGCAACTCTGCGC 240

QY 274 GTGCTGCGCGCGCGGAGCTTACCCGCGCGCGGAGCTTACCGCTGCTGCTGAG 333

DB 241 GTGCTGCGCGCGCGGAGCTTACCCGCGCGCGGAGCTTACCGCTGCTGCTGAG 300

QY 334 CCA 336

DB 301 CCA 303

RESULT 20
 AAD59057 ID AAD59057 standard; cDNA; 303 BP.

AC AAD59057;

DT 18-DEC-2003 (first entry)

DE Human DED4 DED CDNA.

XX Human; death Domain; DD; death effector domain; DED; cell proliferation;

XX Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;

XX neurotrophic factor receptor-interacting death domain; cell adhesion;

XX vasculopathy; microbial infection; inflammation; allograft rejection; CTDD;

XX cell stress responses; benign prostatic hypertrophy; antibacterial; NIDD;

XX apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;

XX neoplasia; restenosis; immunosuppressive; antibody therapy; cytoskeletal;

XX keloid; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..303

FT /tag= a

FT /product= "Human DED4 DED protein"

FT /note= "No start and stop codon"

FT /partial

PM US2003049702-A1.

PD 13-MAR-2003.

PF 15-NOV-2001; 2001US-00001254.

PR 17-NOV-2000; 2000US-00715893.

PR 17-NOV-2000; 2000US-0367360P.

PR 29-JUN-2001; 2001US-0301889P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (PAWL/) PAWLOWSKI K.

XX (FIOR/) FIORENTINO L.

XX (LEES/) LEES S H.

XX (ROTH/) ROTH W.

XX (STEN/) STENNER-LIEWEN F.

XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;

XX Steiner-Liwen F;

XX WPI; 2002-500222/53.

XX P-PSDB; AAE38898.

XX

XX New polypeptide comprising a death domain or death effector domain,

XX useful for discovery of drugs that suppress infection, inflammation,

XX allergy, sepsis, autoimmunity, allograft rejection and other diseases.

XX Claim 18; Page 34-35; 99p; English.

XX The present invention provides novel death Domain (DD) and death effector

XX domain (DED) proteins and nucleic acids encoding them. The invention also

XX provides death domain containing protein such as Chlamydia trachomatis

XX death domain containing protein (CTDD) DD and neural growth factor

XX receptor-interacting death domain (NIDD) DD. The invention is useful for

XX identifying a binding agent (e.g. protein or drug) that binds a DD, DED

XX or NB-ARC domain from DAP3, IRAK4, CTDD, DED4 or NIDD with a candidate

XX binding agent and identifying an effective agent (e.g. protein or drug)

XX that modulates the association of a DD, DED or NB-ARC domain with protein

XX that binds the DD, DED or NB-ARC domain. The invention is also useful for

XX cell proliferation, cell stress responses, responses to microbial

XX infection and B cell immunoglobulin class switching. DEDs and NB-ARC

XX domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are

XX useful for discovery of drugs that suppress infection, autoimmunity,

XX inflammation, allergy, allograft rejection, sepsis and other diseases.

XX DD, DED or NB-ARC domain proteins are used to treat infection, allergy,

XX autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte

XX hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,

XX inflammatory hyperplasia and smooth muscle cell proliferation in arteries

XX following balloon angioplasty (restenosis). The invention is also used in

XX antibody therapy and gene therapy. The present sequence is human DED4 DED

XX CDNA. The DEDD gene is located on chromosome 1

XX

SO Sequence 303 BP; 41 A; 96 C; 115 G; 51 T; 0 U; 0 Other;

Query Match 27.7%; Score 252; DB 6; Length 303;

Best Local Similarity 99.7%; Pred. No. 3.6e-105;

Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 34 TGGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGTGCTTACCGTATGTTGAG 93

DB 1 TGGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGTGCTTACCGTATGTTGAG 60

PF 17-JUL-2001; 2001WO-IL000653.
XX
XX 18-JUL-2000; 2000IL-00137345.
PR 15-DEC-2000; 2000IL-00140354.
XX
XX (COMP-) COMPUGEN LTD.
XX
XX Mintz L, Freilich S, Bernstein J;
PI
DR WPI: 2002-155037/20.
DR P-PSDB; ABB06039.
PT One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
XX
PS Claim 1; Page 77; 290pp; English.
XX
XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antineumatic, antirheumatic, antipneumatic, ophthalmological, virucide,
CC vasotropic, antiatherosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antifertility, cardiovascular,
CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, cerebroprotective,
CC antidepressant, gastrointestinal, neurolipic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC caracerts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive
XX
SQ Sequence 1106 BP; 182 A; 340 C; 385 G; 189 T; 0 U; 10 Other;
Query Match 20.6%; Score 187; DB 6; Length 1106;
Best Local Similarity 100.0%; Pred. No. 2e-75;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 687 GGACGTGTTGGGAGGCGACCGAGTGTGCTGCTCAAGGAGCCTGGGCTGTGGTTG 746
Db 786 GGACGTGTTGGGAGGCGACCGAGTGTGCTGCTCAAGGAGCCTGGGCTGTGGTTG 845
QY 747 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGCGACTAGTGG 806
Db 846 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGCGACTAGTGG 905
QY 807 CGCCCTGCTGCAAGCGCTTGGGGCGGTTCCTGTAAGAGCCTTGGAGAGGCTGTGG 866
Db 906 CGCCCTGCTGCAAGCGCTTGGGGCGGTTCCTGTAAGAGCCTTGGAGAGGCTGTGG 965
QY 867 CCGGGAG 873
Db 966 CCGGGAG 972

RESULT 23
ID AAF93937 standard; DNA; 603 BP.
XX
XX AAF93937;
XX
XX 23-MAY-2001 (first entry)
XX
XX Primer specific for DNA encoding secretory/membrane protein SEQ ID 371.
XX
XX Human; secretory protein; membrane protein; vaccine; gene therapy;
KW rheumatoid arthritis; diabetes; PCR primer; ss.

XX
XX Synthetic.
OS
XX
XX EPI067182-A2.
PN
XX
XX 10-JAN-2001.
PD
XX
XX 07-JUL-2000; 2000EP-00114090.
PF
XX
XX 08-JUL-1999; 99JP-00194179.
PR 11-JAN-2000; 2000JP-00118775.
PR 02-MAY-2000; 2000JP-00183766.
XX
XX (HELI-) HELIX RES INST.
PA
PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
XX WPI: 2001-093989/11.
DR
XX
XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
PT gene therapy or as candidate target molecules in drug development.
XX
XX
PS Claim 4; SEQ ID NO 371; 609pp + Sequence Listing; English.
XX
XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC which encode human secretory or membrane proteins represented by AAB86317
CC - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC invention. The invention also includes methods for the production of
CC antibodies directed against the proteins, and cDNA sequences, which can
CC be used in vaccines. The polynucleotide sequences can be used in gene
CC therapy. The polynucleotide sequences and the proteins they encode may be
CC used in the prevention, treatment and diagnosis of diseases associated
CC with inappropriate secretory protein/membrane protein expression. The
CC nucleic acids and complementary sequences may also be used as DNA probes
CC in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC and quantitate the presence of similar nucleic acid sequences in samples.
CC They may also be used to study the expression and function of secretory
CC proteins/membrane polypeptides and their role in metabolism. The
CC polypeptides may be used as antigens in the production of antibodies
CC against them and in assays to identify modulators (agonists and
CC antagonists) of expression and activity. The antibodies and antagonists
CC may also be used as therapeutic agents to down regulate expression and
CC activity. The antibodies may also be used as diagnostic agents for
CC detecting the presence of the polypeptides in samples (e.g. by enzyme
CC linked immunosorbent assay (ELISA). Examples of diseases which may be
CC treated include rheumatoid arthritis and diabetes
XX
SQ Sequence 603 BP; 114 A; 189 C; 187 G; 103 T; 0 U; 10 Other;
Query Match 19.5%; Score 177; DB 5; Length 603;
Best Local Similarity 100.0%; Pred. No. 8e-71;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCTGGAGAGATGAGTGCCTGACTAC 60
Db 123 ATGGCGCTATCCGGGTCGACCCCGCCGCTGCTGGAGAGATGAGTGCCTGACTAC 182
QY 61 TACGGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGGCAATGACCGAGTGC 120
Db 183 TACGGGATGCTGTGCTTCAACCGTATGTTGAGAGTGTGGCGGGCAATGACCGAGTGC 242
QY 121 GAGCTGAGAGCTCCGGGCTTTCGCTGAGAGAGCTCTGGCGCGCGGAGAGGCTTA 177
Db 243 GAGCTGAGAGCTCCGGGCTTTCGCTGAGAGAGCTCTGGCGCGCGGAGAGGCTTA 299

RESULT 24
ID ACH41209 standard; cDNA; 451 BP.
XX
XX ACH41209;
XX

DT 13-OCT-2003 (first entry)
 XX Human foetal brain CDNA #2576.
 XX
 XX
 KM Human; ss: sequencing by hybridisation; SBH: expressed sequence tag; EST;
 KM genome mapping; biodiversity; genetic disorder.
 XX
 OS Homo sapiens.
 PN US2003073623-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 30-JUL-2001; 2001US-00918995.
 XX
 PR 30-JUL-2001; 2001US-00918995.
 XX
 XX 30-JUL-2001; 2001US-00918995.
 XX
 XX (DRMA/) DRMANAC R T.
 PA (LABA/) LABAT I.
 PA (STAC/) STACHE-CRAIN B.
 PA (DICK/) DICKSON M C.
 PA (JONE/) JONES L W.
 XX
 PI Dmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 XX
 XX WPI; 2003-615964/58.
 XX
 PT New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 XX
 PS Claim 1; SEQ ID NO 28421; 44pp; English.
 XX
 PS The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations,
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030073623
 CC
 XX
 SQ Sequence 451 BP; 78 A; 139 C; 147 G; 84 T; 0 U; 3 Other;
 Query March 16.6%; Score 151; DB 9; Length 451;
 Best Local Similarity 100.0%; Pred. No. 6.ee-59;
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 759 CTCGAGCTCTCCATCTGAGCGCCTTCTGGGGGAGCACTGAGTGGCCCTGCTGCA 818
 Db 42 CTCGAGCTCTCCATCTGAGCGCCTTCTGGGGGAGCACTGAGTGGCCCTGCTGCA 101
 QY 819 GGCCTCTGGGGGCGTGTCTGACTGAGAGCCCTGCGAGAGCTGTGGGCGGAGAGCTGT 878
 Db 102 GGCCTCTGGGGGCGTGTCTGACTGAGAGCCCTGCGAGAGCTGTGGGCGGAGAGCTGT 161
 QY 879 TCGCCTGTGCTGCTGATGTGATGAGGCTGAC 909
 Db 162 TCGCCTGTGCTGATGTGATGAGGCTGAC 192

RESULT 25
 ACH25056

ID ACH25056 standard; cDNA; 484 BP.
 XX
 XX ACH25056;
 XX
 XX
 DT 13-OCT-2003 (first entry)
 XX
 XX
 DE Human adult ovary CDNA #3436.
 XX
 XX
 KM Human; ss: sequencing by hybridisation; SBH: expressed sequence tag; EST;
 KM genome mapping; biodiversity; genetic disorder.
 XX
 OS Homo sapiens.
 PN US2003073623-A1.
 XX
 PD 17-APR-2003.
 XX
 PF 30-JUL-2001; 2001US-00918995.
 XX
 PR 30-JUL-2001; 2001US-00918995.
 XX
 XX 30-JUL-2001; 2001US-00918995.
 XX
 XX (DRMA/) DRMANAC R T.
 PA (LABA/) LABAT I.
 PA (STAC/) STACHE-CRAIN B.
 PA (DICK/) DICKSON M C.
 PA (JONE/) JONES L W.
 XX
 PI Dmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 XX
 XX WPI; 2003-615964/58.
 XX
 PT New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 XX
 PS Claim 1; SEQ ID NO 12268; 44pp; English.
 XX
 PS The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations,
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030073623
 CC
 XX
 SQ Sequence 484 BP; 83 A; 148 C; 159 G; 93 T; 0 U; 1 Other;
 Query March 10.6%; Score 96; DB 9; Length 484;
 Best Local Similarity 100.0%; Pred. No. 1e-33;
 Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 750 CATCAAGTTTCTAGAGCTCTCTATCTGAGCGCCTTCTGGGGGAGCACTGAGTGGCGC 809
 Db 36 CATCAAGTTTCTAGAGCTCTCTATCTGAGCGCCTTCTGGGGGAGCACTGAGTGGCGC 95
 QY 810 CTCGCTGAGAGCCCTTGGGGGCGTGTCTGACTGTA 845
 Db 96 CTCGCTGAGAGCCCTTGGGGGCGTGTCTGACTGTA 131

RESULT 26
 ADC79263

ID ADC79263 standard; DNA; 25 BP.
XX
AC ADC79263;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human DEBD2 related PCR primer SEQ ID NO:5.
XX
KM human; death effector domains containing DNA-binding protein;
KM DEB-containing DNA-binding protein; DEBD2; cell death; gene therapy;
KM cytosol; cancer; chronic myeloid leukaemia; PCR primer; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003054195-A1.
XX
PD 03-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-JP013371.
XX
PR 20-DEC-2001; 2001JP-00387854.
PR 18-JUL-2002; 2002JP-00209458.
XX
PA (MORG) MORINAGA MILK IND CO LTD.
XX
PI Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX
DR WPI; 2003-569246/53.
XX
PT DNA encoding cell death proteins for treatment of kidney, large intestine
PT and prostate cancers and leukemia.
XX
PS Example; Page 22; 26pp; Japanese.
XX
CC The present invention describes a human death effector domains (DED)
CC containing DNA-binding protein (DEBD) protein, designated DEBD2, that
CC causes cell death. Also described: (1) primer and probe for investigation
CC of the DEBD2 gene; and (2) reagents for gene therapy. DEBD2 has
CC cytosolic activity. DEBD2 can be used in the diagnosis and treatment of
CC cancers of the kidney, large intestine and prostate, and acute and
CC chronic myeloid leukaemia. The present sequence represents a PCR primer
CC which is used in an example from the present invention.
XX
SQ Sequence 25 BP; 5 A; 9 C; 6 G; 5 T; 0 U; 0 Other;
XX
Query Match 2.8%; Score 25; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 321 CCGGCCAGTGTCTCCAGAACGCTAT 345
DB 1 CCGGCCAGTGTCTCCAGAACGCTAT 25
XX
RESULT 27
ADL44878/c
ID ADL44878 standard; DNA; 838 BP.
XX
AC ADL44878;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #18768.
XX
KM Human; ovarian cancer; ds; tumour; cytosol; DNA marker.
XX
OS Homo sapiens.
XX
PN WO200170979-A2.
XX
PD 27-SEP-2001.
XX

PF 21-MAR-2001; 2001WO-US009126.
XX
XX 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 18768; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 838 BP; 165 A; 304 C; 147 G; 208 T; 0 U; 14 Other;
XX
Query Match 2.4%; Score 22; DB 5; Length 838;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 35 GGGAGAGAGATGAGTGGCTGGA 56
DB 381 GGGAGAGAGATGAGTGGCTGGA 360
XX
RESULT 28
ADL39650
ID ADL39650 standard; DNA; 323 BP.
XX
AC ADL39650;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human ovarian cancer DNA marker #13540.
XX

XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
 KW Homo sapiens.
 OS
 XX WO200170979-A2.
 PN
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009126.
 XX
 PR 21-MAR-2000; 2000US-0191031P.
 PR 25-MAY-2000; 2000US-0207124P.
 PR 15-JUN-2000; 2000US-0211940P.
 PR 07-JUL-2000; 2000US-0216820P.
 PR 25-JUL-2000; 2000US-0220661P.
 PR 21-DEC-2000; 2000US-0257672P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI
 PI Lee J, Little J;
 XX
 WPI; 2001-611502/70.
 DR
 XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
 PT cancer cells as compared to their normal non-cancerous ovarian cells are
 PT used to characterize stage, grade, histological type of ovarian cancer.
 XX
 PS Disclosure; SEQ ID NO 13540; 106bp; English.
 XX
 CC The invention relates to nucleic acid markers which are overexpressed in
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-
 CC cancerous) ovarian cells. The invention also relates to polypeptides
 CC encoded by the markers, antibodies that selectively bind to the
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
 CC of developing ovarian cancer involving inhibiting expression of a gene
 CC corresponding to a marker of the invention and a method of treating a
 CC patient afflicted with ovarian cancer comprising providing to cells of
 CC the patient an antisense oligonucleotide complementary to a marker of the
 CC invention. The markers are useful for assessing if a patient is afflicted
 CC with ovarian cancer, which involves comparing the level of expression of
 CC a marker in a patient sample and a normal level of expression of the
 CC marker in a control non-ovarian cancer sample. A difference between the
 CC expression levels indicates ovarian cancer. The level of expression of a
 CC marker corresponds to a secreted protein or to a transcribed
 CC polynucleotide or its portion. The level of expression of the marker is
 CC assessed by detecting the presence in the sample, a protein or protein
 CC fragment corresponding to the marker. The presence of protein or protein
 CC fragment is detected using an antibody that specifically binds with the
 CC protein or protein fragment. Alternatively, the level of expression of
 CC the marker is assessed by detecting the presence of a transcribed
 CC polynucleotide which anneals with the marker or anneals with a portion of
 CC the polynucleotide comprising the marker, under stringent conditions. The
 CC marker is also used for monitoring the progression of ovarian cancer in a
 CC patient which involves detecting expression of the marker in a patient
 CC sample at a first point in time, repeating the method at a subsequent
 CC time and comparing the level of expression. The method is carried out
 CC using an ovarian tissue sample. A composition comprising a marker,
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.
 CC This sequence represents a human ovarian cancer DNA marker of the
 CC invention.
 CC
 XX Sequence 323 BP; 87 A; 87 C; 81 G; 68 T; 0 U; 0 Other;
 SQ
 Query Match 2.3%; Score 21; DB 5; Length 323;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 582 GACCTGTGACATCCGGCTCCG 602
 DB 279 GACCTGTGACATCCGGCTCCG 299

RESULT 29
 AAF27413
 ID AAF27413 standard; DNA; 21 BP.
 XX
 XX AAF27413;
 AC
 XX 24-APR-2001 (first entry)
 DT
 XX Human apoptosis-associated factor NT2RM1000558 PCR primer, SEQ ID NO:10.
 DE
 XX Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
 KW DDD; caspase family cleavage domain; pro-apoptotic; drug screening;
 KW cell proliferation; ischemic disease; chronic viral disease; PCR primer;
 KW ss.
 XX
 XX Homo sapiens.
 OS
 XX WO200104300-A1.
 PN
 PD 18-JAN-2001.
 XX
 PF 06-JUL-2000; 2000WO-JP004516.
 XX
 PR 08-JUL-1999; 99JP-001994179.
 PR 18-OCT-1999; 99US-0159586P.
 XX
 PA (HELI-) HELIX RES INST.
 PI
 PI Oca T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
 XX
 WPI; 2001-138348/14.
 DR
 XX Polynucleotide encoding an apoptosis-associated factor protein with death
 PT effector domain and caspase family-cleavage domain, useful in regulating
 PT diseases with cell proliferation.
 XX
 PS Example 4; Page 32; 53bp; Japanese.
 XX
 CC The invention relates to a novel human apoptosis-associated factor
 CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
 CC effector domain (DED) and a caspase family cleavage domain and is capable
 CC of inducing apoptosis in cells. The invention also relates to nucleic
 CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
 CC (particularly dominant negative variants); vectors and host cells
 CC comprising a nucleic acid which encodes an apoptosis-associated factor
 CC of the invention; the recombinant production of the protein; an antibody
 CC against the protein; and methods of screening for compounds which can
 CC regulate apoptosis. The apoptosis-related factor is useful in regulating
 CC diseases associated with cell proliferation and in screening drug
 CC candidates e.g., for regulating cell proliferation or cell death in
 CC ischaemic diseases and chronic viral diseases. The present sequence
 CC represents a PCR primer used to amplify cDNA encoding the human apoptosis
 CC -associated factor NT2RM1000558
 CC
 XX Sequence 21 BP; 4 A; 3 C; 8 G; 6 T; 0 U; 0 Other;
 SQ
 Query Match 2.2%; Score 20; DB 5; Length 21;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 890 TCAGTGTGATGAGGCTGAC 909
 DB 1 TCAGTGTGATGAGGCTGAC 20
 RESULT 30
 AAF27416/C
 ID AAF27416 standard; DNA; 23 BP.
 XX
 XX AAF27416;
 AC
 XX 24-APR-2001 (first entry)
 DT
 XX

DE Human apoptosis-associated factor NT2RM1000558 PCR primer, SEQ ID NO:13.
XX
XX Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
KM DED; caspase family cleavage domain; pro-apoptotic; drug screening;
KM cell proliferation; ischaemic disease; chronic viral disease; PCR primer;
KM 89.
XX
XX Homo sapiens.
XX
XX WO200104300-A1.
XX
XX 18-JAN-2001.
XX
XX 06-JUL-2000; 2000WO-JP004516.
XX
XX 08-JUL-1999; 99JP-00194179.
XX 18-OCT-1999; 99US-0159586P.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oca T, Isegai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
XX WPI; 2001-138348/14.
XX
XX Polynucleotide encoding an apoptosis-associated factor protein with death
PT effector domain and caspase family-cleavage domain, useful in regulating
PT diseases with cell proliferation.
XX
XX Example 5; Page 33; 53pp; Japanese.
XX
XX The invention relates to a novel human apoptosis-associated factor
CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
CC effector domain (DED) and a caspase family cleavage domain and is capable
CC of inducing apoptosis in cells. The invention also relates to nucleic
CC acids encoding the protein (AAFP27407, AAFP27408); variants of the protein
CC (particularly dominant negative variants); vectors and host cells
CC comprising a nucleic acid which encodes an apoptosis-associated factor
CC of the invention; the recombinant production of the protein; an antibody
CC against the protein; and methods of screening for compounds which can
CC regulate apoptosis. The apoptosis-related factor is useful in regulating
CC diseases associated with cell proliferation and in screening drug
CC candidates e.g., for regulating cell proliferation or cell death in
CC ischaemic diseases and chronic viral diseases. The present sequence
CC represents a PCR primer used to amplify cDNA encoding the human apoptosis
CC -associated factor NT2RM1000558
XX
XX Sequence 23 BP; 6 A; 9 C; 3 G; 5 T; 0 U; 0 Other;
SQ
Query Match 2.2%; Score 20; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 890 TCAGTGTGATGAGGCTGAC 909
Db 23 TCAGTGTGATGAGGCTGAC 4
RESULT 31
ACCS9397
ID ACCS9397 standard; DNA; 2610 BP.
XX
XX ACCS9397;
XX
XX 28-AUG-2003 (first entry)
XX
XX Microbial resistance gene NdVb coding sequence.
XX
XX Biofilm; microbial resistance; gene; ds.
XX
XX Unidentified.
XX
XX WO2003041483-A2.
XX
XX

PD 22-MAY-2003.
XX
XX 18-SEP-2002; 2002WO-US029565.
XX
XX 18-SEP-2001; 2001US-0323241P.
XX
XX (DART-) DARTMOUTH COLLEGE.
XX
XX O'toole GA, Mah T;
XX
XX WPI; 2003-468567/44.
XX
XX P-PSDB; ABP59932.
XX
XX Identifying modulators of microbial resistance of organisms in biofilms,
PT e.g. inhibitor of biofilm formation, by employing expression controls, or
PT efflux pumps containing polypeptides, of genes associated with biofilm
PT resistance.
XX
XX Disclosure; Fig 6; 102pp; English.
XX
XX The present invention relates to a method of identifying a compound
CC capable of altering the sensitivity of a microorganism to an
CC antimicrobial agent by employing efflux pumps comprising polypeptides
CC encoded by the following genes: PA1874, PA1442, PA2389, PA1876, PA143,
CC PA2390 or PA1163. The method is useful for identifying modulators of
CC microbial resistance of an organism in a biofilm. The methods are also
CC useful for identifying genes that encode proteins that play a role in
CC biofilm resistance. The method is particularly useful for screening
CC compounds or discovering compositions that will inhibit biofilm formation
CC and overcome their resistance mechanisms. These methods are particularly
CC useful in medical, industrial or natural settings, where formation of
CC biofilms can have serious negative consequences and result in high costs
CC both in human health and economic terms. The present sequence is a
CC coding/control sequence shown in the exemplification of the invention
XX
XX Sequence 2610 BP; 391 A; 952 C; 839 G; 428 T; 0 U; 0 Other;
SQ
Query Match 2.2%; Score 20; DB 9; Length 2610;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 308 CGCGCAAGCGCGCGGCCA 327
Db 1208 CGCGCAAGCGCGCGGCCA 1227
RESULT 32
ABD04389/C
ID ABD04389 standard; DNA; 2673 BP.
XX
XX ABD04389;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polynucleotide #2993.
XX
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
XX antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US651795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI

CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptide or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX SQ Sequence 2814 BP; 422 A; 1026 C; 908 G; 458 T; 0 U; 0 Other;
Query Match 2.2%; Score 20; DB 11; Length 2814;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 CGCGCAAGCGCGCGGCCA 327
Db 1371 CGCGCAAGCGCGCGGCCA 1390
|||||
|||||

RESULT 35
ADE78561
ID ADE78561 standard; DNA; 360 BP.
XX
AC ADE78561;
XX
DT 29-JAN-2004 (first entry)
XX
DE 35 amino acid human BAG3 protein partially encoding DNA.
XX
XX BAG3; vasotrophic; antiparkinsonian; neuroprotective; cytosolic;
KW vulnery; antitumor; immunosuppressive; antidiabetic; cardiant;
KW cerebroprotective; anti-HIV; human; gene; ds.
XX
XX Homo sapiens.
OS
XX
XX EP1323733-A1.
PN
XX
XX 02-JUL-2003.
PD
XX
XX 28-DEC-2001; 2001EP-00830834.
PF
XX
XX 28-DEC-2001; 2001EP-00830834.
PR
XX
XX (LEON/) LEONE A.
PA (TURC/) TURCO M C.
XX
XX Leone A, Turco MC;
PI
XX
XX WPI; 2003-723295/69.
DR P-PSDB; ADE78562.
XX
XX
PT New isolated BAG3 protein, useful for treating Parkinson's disease,
PT amyotrophic lateral sclerosis, skeletal muscle disorders, tissue repair
PT or wound healing, and treatment of surgical incisions.
XX
XX
XX Claim 1; SEQ ID NO 10; 42pp; English.

CC The invention relates to a novel isolated polypeptide having a fully
CC defined BAG3 protein sequence of 575 amino acids, 35 amino acids, 395
CC amino acids or 67 amino acids and relative fragments having a fully
CC defined amino acid sequence of 15 amino acids as given in the
CC specification. The BAG3 polypeptides have the following activities:
CC vasotrophic, antiparkinsonian, neuroprotective, cytosolic, vulnery,
CC antitumor, immunosuppressive, antidiabetic, cardiant, cerebroprotective,
CC and anti-HIV. The BAG3 polypeptides and fragments; the encoding nucleic
CC acids, complements and their fragments; and a BAG3 carrier are useful for
CC modulating apoptosis in primary cells and treating diseases chosen from
CC acute or chronic tissue damages, such as heart, kidney, brain or other
CC organ ischaemia, HIV-related damage of brain or other tissues, skeletal
CC muscle disorders, transplantation rejection, chronic degenerative

CC disorders such as Parkinson's disease, amyotrophic lateral sclerosis and
CC others, and neoplastic, autoimmune and other diseases involving excessive
CC or defective apoptosis, tissue repair or wound healing, treatment of
CC surgical incisions, and ulcers, such as stomach or diabetic ulcers. This
CC polynucleotide sequence represents the DNA partially encoding the 35
CC amino acid human BAG3 protein of the invention.

XX SQ Sequence 360 BP; 58 A; 146 C; 101 G; 55 T; 0 U; 0 Other;
Query Match 2.1%; Score 19; DB 10; Length 360;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 508 CGCGCGCGAGAGGGGCC 526
Db 207 CGCGCGCGAGAGGGGCC 225
|||||
|||||

RESULT 36
ADQ55265
ID ADQ55265 standard; DNA; 579 BP.
XX
AC ADQ55265;
XX
DT 21-OCT-2004 (first entry)
XX
XX Novel canine microarray-related DNA sequence SeqID6567.
XX
XX canine microarray; drug screening; toxicity assay;
KW environmental pollutant; cellular response; gene expression profile;
KW toxic response; liver necrosis; fatty liver disease;
KW protein adduct formation; hepatitis; dog; ds.
XX
XX
XX Canis familiaris.
OS
XX
XX WO2004063324-A2.
PN
XX
XX 29-JUL-2004.
PD
XX
XX 05-MAY-2003; 2003WO-US013853.
PF
XX
XX 03-MAY-2002; 2002US-0377240P.
PR
XX
XX (GENE-) GENE LOGIC INC.
PA (PRIZ) PRIZER PROD INC.
XX
XX Diggins JC, Porter M, Wei T;
PI
XX
XX WPI; 2004-561890/54.
DR
XX
XX
XX
PT New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
XX
XX Claim 1; SEQ ID NO 6567; 41pp; English.

CC This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the

CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.

XX Sequence 579 BP, 107 A; 170 C; 163 G; 127 T; 0 U; 12 Other;

Query Match 2.1%; Score 19; DB 13; Length 579;

Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 808 GCCCTGCTGCAGGCCCTGC 826
Db 101 GCCCTGCTGCAGGCCCTGC 119

RESULT 37

AAH52281 ID AAH52281 standard; cDNA; 606 BP.

XX AAH52281;

DT 10-SEP-2001 (first entry)

DE Human AFP protein encoding cDNA sequence SEQ ID NO:377.

KM Human; secreted protein; secretion; bacterial cell; fungal cell;

KW eukaryotic cell; fusion protein; maltose binding protein;

KX immunoglobulin constant region; polyhistidine tag; ss.

OS Homo sapiens.

PN WO200129221-A2.

PD 26-APR-2001.

PF 20-OCT-2000; 2000WO-US029952.

PR 20-OCT-1999; 99US-0160712P.

PA (ZYMO) ZYMOGENETICS INC.

PI Conklin DC, Yee DP;

DR WPI; 2001-300340/31.

DR P-PSDB; AAG81430.

XX Isolated polypeptide for directing secretion of proteins of interest from
PT a host cell including, e.g., bacteria, includes contiguous amino acid
PT residues of polypeptide with specified amino acids.

PS Claim 9; Page 585-586; 617pp; English.

CC AAH52093 to AAH52303 encode the human secreted proteins given in AAG81242

CC to AAG81453. The secreted proteins can be used for directing the

CC secretion of proteins of interest from a host cell including bacteria,

CC fungal cells, and cultured higher eukaryotic cells. The present invention

CC also describes fusion proteins, where a secreted protein of the invention

CC is operably linked via a peptide bond or peptide linker to a second

CC protein selected from the group consisting of maltose binding protein, an

CC immunoglobulin constant region, a polyhistidine tag and a peptide given

CC in AAG81453

XX Sequence 606 BP; 93 A; 198 C; 205 G; 110 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 5; Length 606;
Best Local Similarity 100.0%; Pred. No. 2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 25 GCCCGGCTGAGAGAGAG 43
Db 391 GCCCGGCTGAGAGAGAG 409

RESULT 38

ABK70016 ID ABK70016 standard; DNA; 943 BP.

XX ABK70016;

DT 15-JUL-2002 (first entry)

DE cDNA encoding human pro peptide #56.

KM Human; ss; gene; PRO; secreted protein; transmembrane protein;

KW genetic disorder; tumour; cancer.

OS Homo sapiens.

PN WO200224888-A2.

PD 28-MAR-2002.

PF 29-AUG-2001; 2001WO-US027099.

PR 01-SEP-2000; 2000US-0229886P.

PR 05-SEP-2000; 2000US-0230621P.

PR 22-SEP-2000; 2000US-0235147P.

PR 10-NOV-2000; 2000WO-US030873.

PR 12-JAN-2001; 2001US-0261878P.

PR 16-JAN-2001; 2001US-0261910P.

PR 16-JAN-2001; 2001US-0261939P.

PR 16-JAN-2001; 2001US-0262150P.

PR 25-JAN-2001; 2001US-0264395P.

PR 02-FEB-2001; 2001US-0266421P.

PR 09-FEB-2001; 2001US-0267623P.

PR 28-FEB-2001; 2001WO-US006520.

PR 09-MAR-2001; 2001US-0274399P.

PR 03-APR-2001; 2001US-0280982P.

PR 04-APR-2001; 2001US-0282129P.

PR 04-APR-2001; 2001US-0282199P.

PR 09-MAY-2001; 2001US-0290589P.

PR 25-MAY-2001; 2001WO-US017092.

PR 01-JUN-2001; 2001WO-US017800.

PR 20-JUN-2001; 2001WO-US019692.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-JUL-2001; 2001WO-US021735.

PA (GERTH) GENENTECH INC.

PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;

PI Gunney AL, Smith V, Stephan J, Watanabe CK, Wood WJ, Zhang Z;

PI Fong S;

DR WPI; 2002-362426/39.

DR P-PSDB; ABG34085.

XX New PRO polypeptides and polynucleotides encoding the polypeptides,

PT useful in gene therapy, chromosome identification, tissue typing, or for
PT genetic analysis of individuals with genetic disorders.
XX Claim 2; Fig 111; 218pp; English.
XX This invention relates to the cDNA and protein sequences of novel
CC secreted and transmembrane polypeptides PRO polypeptides. The invention
CC also comprises a method for producing the proteins of the invention by
CC recombinant means and antibodies specific for the protein of the
CC invention. The antibody may be used for detecting the PRO proteins of the

CC invention and may be used to modify their activity. polynucleotides may
CC be used as hybridisation probes for a cDNA library to isolate the full-
CC length PRO cDNA or to isolate other cDNAs, to construct hybridisation
CC probes for mapping the gene which encodes that PRO and for genetic
CC analysis of individuals with genetic disorders, in assays to identify
CC other proteins or molecules involved in binding reaction, to generate
CC transgenic animals or knock-out animals which in turn are useful in the
CC development and screening of therapeutically useful reagents, for
CC chromosome identification, and tissue typing. The PRO polypeptides are
CC useful in gene therapy, and as molecular weight markers for protein
CC electrophoresis purposes. The sequences may also be used to detect
CC overexpression on PRO polypeptides in cancerous tumours and for screening
CC for differentially expressed genes using microarray technology. The
CC present sequence represents a cDNA encoding a human PRO protein of the
CC invention
XX

SO Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
Query Match 2.1%; Score 19; DB 6; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 GCGCCGTCGTGGAGAGG 43
DB 437 GCGCCGTCGTGGAGAGG 455

RESULT 39
ADA01379
ID ADA01379 standard; cDNA; 943 BP.
XX
AC ADA01379;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human PRO polynucleotide #56.
XX
XX Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
KM tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
KM adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
KM microvascular endothelial cell; endothelial cell tube formation;
KM sports-related joint problem; articular cartilage defect; osteoarthritis;
KM rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
XX
OS Homo sapiens.
XX
PN US2003068779-A1.
XX
PD 10-APR-2003.
XX
PF 16-SEP-2002; 2002US-00245107.
XX
PR 09-MAY-2001; 2001US-0280589P.
PR 29-AUG-2001; 2001WO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
XX (GETH) GENENTECH INC.
XX
PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX
DR WPI; 2003-625484/59.
XX P-PSDB; ADA01380.
XX
XX Novel isolated PRO1313, PRO20080 or PRO21383 polypeptide useful for
PT stimulating proliferation of human microvascular endothelial cells, and
PT PRO6018 polypeptide useful for stimulating proliferation of chondrocyte
PT cells.
XX
PS Claim 2; Fig 111; 307pp; English.
XX
XX The invention relates to isolated human PRO polypeptides (secreted and

CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
CC colon, breast, prostate, rectal, cervical and liver tumours). The
CC polynucleotides are useful in molecular biology, including uses as
CC hybridisation probes, in chromosome and gene mapping, in generating
CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
CC be used in preparing PRO polypeptides by recombinant techniques and in
CC generating either transgenic animals or knock-out animals which are
CC useful in the development and screening of therapeutically useful
CC reagents. The PRO polypeptides or antibodies are used in preparing a
CC medicament for treating a condition responsive to the polypeptides or
CC antibodies, such as tumours, for stimulating and inhibiting proliferation
CC of human microvascular endothelial cells, for inducing endothelial cell
CC tube formation and for treating sports-related joint problems, articular
CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
CC represents a human PRO polynucleotide of the invention.
XX

SO Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
Query Match 2.1%; Score 19; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 25 GCGCCGTCGTGGAGAGG 43
DB 437 GCGCCGTCGTGGAGAGG 455

RESULT 40
ADA43808
ID ADA43808 standard; cDNA; 943 BP.
XX
AC ADA43808;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human cDNA encoding secreted/transmembrane polypeptide PRO21434.
XX
XX ss; gene; human; PRO; secreted protein; transmembrane protein;
KM endothelial cell tube formation; chondrocyte cell differentiation;
KM microvascular endothelial cell; tumour; lung tumour; colon tumour;
KM breast tumour; prostate tumour; rectal tumour; kidney tumour;
KM liver tumour; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN US2003064474-A1.
XX
PD 03-APR-2003.
XX
PF 16-SEP-2002; 2002US-00245859.
XX
PR 29-AUG-2001; 2001WO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
XX (GETH) GENENTECH INC.
XX
PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX
DR WPI; 2003-605867/57.
XX P-PSDB; ADA43809.
XX
XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
PT PRO21383, useful in molecular biology, chromosome and gene mapping, in
PT generating antisense RNA and DNA, and in gene therapy.
XX
XX Claim 2; Fig 111; 308pp; English.
XX
XX

XX The invention relates to an isolated secreted/transmembrane (PRO)
 CC polypeptide, having at least 80% sequence identity to a sequence selected
 CC from any one of the 57 amino acid sequences given in specification, or to
 CC a sequence encoded by a nucleic acid molecule selected from any one of
 CC the nucleic acids deposited under any of the ATCC accession numbers given
 CC in specification, or a sequence having at least 80% identity to PRO
 CC lacking its associated signal peptide, an extracellular domain of PRO
 CC with or without its associated signal peptide. Also included are vectors,
 CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 CC administering PRO281, PRO1560, PRO189, PRO499, PRO6308, PRO6000,
 CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
 CC an oligonucleotide probe derived from any one of the above nucleotide
 CC sequences. PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
 CC and PRO21383 polypeptides are useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
 CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a
 CC medicament for treating a condition responsive to PRO polypeptide. The
 CC oligonucleotide probes are useful for isolating genomic and cDNA
 CC nucleotide sequences, for measuring or detecting the expression of an
 CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
 CC hybridisation probe, in chromosome and gene mapping, in the generation of
 CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.

SQ Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 9; Length 943;

Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 GCGCCGCTGCTGGAGAGG 43

437 GCGCCGCTGCTGGAGAGG 455

RESULT 41

ADA43576

ADA43576 standard; cDNA; 943 BP.

ADA43576;

20-NOV-2003 (first entry)

Human cDNA encoding secreted/transmembrane polypeptide PRO21434.

ss; gene; human; PRO; secreted protein; transmembrane protein;

endothelial cell tube formation; chondrocyte cell differentiation;

microvascular endothelial cell; tumour; lung tumour; colon tumour;

breast tumour; prostate tumour; rectal tumour; kidney tumour;

liver tumour; cytostatic; vaccine.

Homo sapiens.

US2003073196-A1.

17-APR-2003.

18-SEP-2002; 2002US-00246210.

04-APR-2001; 2001US-0282199.

29-AUG-2001; 2001WO-US027099.

18-JUL-2002; 2002US-00197942.

PA (GETH) GENENTECH INC.
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX WPI; 2003-743814/70.
 DR P-PSDB; ADA43577.
 XX New secreted and transmembrane PRO polypeptides, e.g. PRO20080 or
 PT PRO21383 useful for stimulating the proliferation or differentiation of
 PT chondrocyte cells and detecting the presence of a tumor in a mammal.
 XX Claim 2; Fig 111; 307pp; English.

XX The invention relates to an isolated secreted/transmembrane (PRO)
 CC polypeptide, having at least 80% sequence identity to a sequence selected
 CC from any one of the 57 amino acid sequences given in specification, or to
 CC a sequence encoded by a nucleic acid molecule selected from any one of
 CC the nucleic acids deposited under any of the ATCC accession numbers given
 CC in specification, or a sequence having at least 80% identity to PRO
 CC lacking its associated signal peptide, an extracellular domain of PRO
 CC with or without its associated signal peptide. Also included are vectors,
 CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
 CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
 CC administering PRO281, PRO1560, PRO189, PRO499, PRO6308, PRO6000,
 CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
 CC an oligonucleotide probe derived from any one of the above nucleotide
 CC sequences. PRO6018 polypeptide is useful for stimulating the
 CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
 CC and PRO21383 polypeptides are useful for stimulating the proliferation of
 CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
 CC polypeptides are useful for inhibiting the proliferation of human
 CC microvascular endothelial cells. PRO polypeptides are useful for
 CC detecting the presence of tumour in a mammal, including tumours of lung,
 CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
 CC PRO189, PRO499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
 CC PRO34274 polypeptides are useful for inducing endothelial cell tube
 CC formation. PRO or the antibody are useful in the preparation of a
 CC medicament for treating a condition responsive to PRO polypeptide. The
 CC oligonucleotide probes are useful for isolating genomic and cDNA
 CC nucleotide sequences, for measuring or detecting the expression of an
 CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
 CC hybridisation probe, in chromosome and gene mapping, in the generation of
 CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
 CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
 CC present sequence encodes a PRO protein.

SQ Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 9; Length 943;

Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 GCGCCGCTGCTGGAGAGG 43

437 GCGCCGCTGCTGGAGAGG 455

RESULT 42

ADA01251

ADA01251 standard; cDNA; 943 BP.

ADA01251;

06-NOV-2003 (first entry)

Human PRO polynucleotide #56.

Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;

tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;

cancer; lung; colon; breast; prostate; rectum; kidney; liver;

microvascular endothelial cell; endothelial cell tube formation.

XX Homo sapiens.
OS
XX US2003068782-A1.
XX
XX 10-APR-2003.
XX
XX 16-SEP-2002; 2002US-00245851.
XX
XX 27-APR-1999; 99US-0131271P.
PR 29-OCT-1999; 99US-0162506P.
PR 02-DEC-1999; 99MO-US028551.
PR 29-AUG-2001; 2001MO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX
XX WPI: 2003-625487/59.
DR P-PSDB; ADA01252.
XX
XX Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
PT preparation of a medicament for treating a condition responsive to PRO
PT polypeptide, and as therapeutic agents e.g. vaccines.
XX
XX Claim 2; Fig 111; 308pp; English.
XX
XX The invention relates to isolated human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. lung, colon, breast,
CC prostate, rectal, kidney and liver tumours). The polynucleotides are
CC useful in molecular biology, including uses as hybridisation probes, in
CC chromosome and gene mapping, in generating antisense RNA and DNA and in
CC gene therapy. The polynucleotides may also be used in preparing PRO
CC polypeptides by recombinant techniques and in generating either
CC transgenic animals or knock-out animals which are useful in the
CC development and screening of therapeutically useful reagents. The PRO
CC polypeptides or antibodies are used in preparing a medicament for
CC treating a condition responsive to the polypeptides or antibodies, such
CC as tumours, for stimulating and inhibiting proliferation of human
CC microvascular endothelial cells and for inducing endothelial cell tube
CC formation. This sequence represents a human PRO polynucleotide of the
CC invention.
XX
XX Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
SQ
Query Match 2.1%; Score 19; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 GCCCGTGTGGAGAGG 43
Db 437 GCCCGTGTGGAGAGG 455
RESULT 43
ADA01135
ID ADA01135 standard; CDNA, 943 BP.
XX
XX ADA01135;
AC
XX 06-NOV-2003 (first entry)
DT
XX Human cDNA encoding secreted/transmembrane polypeptide PRO21434.
DB
XX 88; gene; human; PRO; secreted protein; transmembrane protein;

KW endothelial cell tube formation; chondrocyte cell differentiation;
KW microvascular endothelial cell; tumour; lung tumour; colon tumour;
KW breast tumour; prostate tumour; rectal tumour; kidney tumour;
KW liver tumour; cytosstatic; vaccine.
XX
XX Homo sapiens.
OS
XX US2003068780-A1.
XX
XX 10-APR-2003.
XX
XX 16-SEP-2002; 2002US-00245143.
XX
XX 02-AUG-2000; 2000US-0222695P.
PR 20-JUN-2001; 2001MO-US019692.
PR 29-AUG-2001; 2001MO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX
XX WPI: 2003-625485/59.
DR P-PSDB; ADA01136.
XX
XX Novel isolated PRO polypeptides e.g. PRO281 and PRO1560, useful in the
PT preparation of a medicament for treating a condition responsive to PRO
PT polypeptide, and as therapeutic agents e.g. vaccines.
XX
XX Claim 2; Fig 111; 307pp; English.
XX
XX The invention relates to an isolated secreted/transmembrane (PRO)
CC polypeptide, having at least 80% sequence identity to a sequence selected
CC from any one of the 57 amino acid sequences given in specification, or to
CC a sequence encoded by a nucleic acid molecule selected from any one of
CC the nucleic acids deposited under any of the ATCC accession numbers given
CC in specification, or a sequence having at least 80% identity to PRO
CC lacking its associated signal peptide, an extracellular domain of PRO
CC with or without its associated signal peptide. Also included are vectors,
CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
CC an oligonucleotide probe derived from any one of the above nucleotide
CC sequences. PRO6018 polypeptide is useful for stimulating the
CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
CC and PRO21383 polypeptides are useful for stimulating the proliferation of
CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
CC polypeptides are useful for inhibiting the proliferation of human
CC microvascular endothelial cells. PRO polypeptides are useful for
CC detecting the presence of tumour in a mammal, including tumours of lung,
CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO20933 and
CC PRO34274 polypeptides are useful for inducing endothelial cell tube
CC formation. PRO or the antibody are useful in the preparation of a
CC medicament for treating a condition responsive to PRO polypeptide. The
CC oligonucleotide probes are useful for isolating genomic and cDNA
CC nucleotide sequences, for measuring or detecting the expression of an
CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.
XX
XX Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
SQ
Query Match 2.1%; Score 19; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 GCCCGTGTGGAGAGG 43

Db 437 GCCCGGTGCTGGAGGAGG 455

RESULT 44
ADA43692
ID ADA43692 standard; cDNA; 943 BP.
XX ADA43692;
AC
XX
XX
XX 20-NOV-2003 (first entry)
XX
XX
XX Human cDNA encoding secreted/transmembrane polypeptide PRO21434.
XX
XX ss; gene; human; PRO; secreted protein; transmembrane protein;
XX endothelial cell tube formation; chondrocyte cell differentiation;
XX microvascular endothelial cell; tumour; lung tumour; colon tumour;
XX breast tumour; prostate tumour; rectal tumour; kidney tumour;
XX liver tumour; cytostatic; vaccine.
XX
XX Homo sapiens.
XX
XX US2003073190-A1.
XX
XX 17-APR-2003.
XX
XX 09-SEP-2002; 2002US-00238283.
XX
XX 01-JUL-1998; 98US-0091358P.
XX 02-JUN-1999; 99MO-US012252.
XX 20-JUL-1999; 99US-0144758P.
XX 28-JUL-1999; 99US-0146222P.
XX 25-AUG-1999; 99US-00380137.
XX 30-MAR-2000; 2000MO-US008439.
XX 02-JUN-2000; 2000MO-US015264.
XX 29-AUG-2001; 2001MO-US027099.
XX 18-JUL-2002; 2002US-00197942.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Eaton DJ, Filvaroff E, Goddard A, Grimaldi JC;
XX Gunney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
XX Fong S;
XX
XX WPI; 2003-585304/55.
XX P-PSDB; ADA43693.
XX
XX New isolated nucleic acid encoding a PRO polypeptide, e.g. PRO20080 or
XX PRO21383, useful in molecular biology, chromosome and gene mapping, in
XX generating antisense RNA and DNA, and in gene therapy.
XX
XX Claim 2; Fig 111; 352pp; English.

XX
XX The invention relates to an isolated secreted/transmembrane (PRO)
XX polypeptide, having at least 80% sequence identity to a sequence selected
XX from any one of the 57 amino acid sequences given in specification, or to
XX a sequence encoded by a nucleic acid molecule selected from any one of
XX the nucleic acids deposited under any of the ATCC accession numbers given
XX in specification, or a sequence having at least 80% identity to PRO
XX lacking its associated signal peptide, an extracellular domain of PRO
XX with or without its associated signal peptide. Also included are vectors,
XX transformed host cells, anti-PRO antibodies, the nucleic acids encoding
XX PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
XX administering PRO281, PRO1560, PRO4499, PRO6308, PRO6000,
XX PRO10975, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
XX an oligonucleotide probe derived from any one of the above nucleotide
XX sequences. PRO6018 polypeptide is useful for stimulating the
XX proliferation or differentiation of chondrocyte cells. PRO313, PRO20080
XX and PRO21383 polypeptides are useful for stimulating the proliferation of
XX human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
XX polypeptides are useful for inhibiting the proliferation of human
XX microvascular endothelial cells. PRO polypeptides are useful for
XX detecting the presence of tumour in a mammal, including tumours of lung,

CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
CC PRO34274 polypeptides are useful for inducing endothelial cell tube
CC formation. PRO or the antibody are useful in the preparation of a
CC medicament for treating a condition responsive to PRO polypeptide. The
CC oligonucleotide probes are useful for isolating genomic and cDNA
CC nucleotide sequences, for measuring or detecting the expression of an
CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and
CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.

XX
XX SQ Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
XX
XX Query Match 2.1%; Score 19; DB 9; Length 943;
XX Best Local Similarity 100.0%; Pred. No. 1.9e+02;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GCCCGGTGCTGGAGGAGG 43
Db 437 GCCCGGTGCTGGAGGAGG 455

RESULT 45
ADA06954
ID ADA06954 standard; cDNA; 943 BP.
XX
XX ADA06954;
AC
XX
XX 06-NOV-2003 (first entry)
XX
XX Human PRO polynucleotide #56.
XX
XX
XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
XX tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
XX adrenal; lung; colon; breast; prostate; rectum; cervix; liver;
XX microvascular endothelial cell; endothelial cell tube formation;
XX bone disorder; cartilage disorder; sports injury; proteoglycan;
XX cartilage; sports-related joint problem; articular cartilage defect;
XX osteoarthritis; rheumatoid arthritis; haemoglobin-associated disorder;
XX thalassemia; immune system cell infiltration; cancer; vulneryary;
XX antianemic; osteopathic; antirheumatic; antiarthritic.
XX
XX Homo sapiens.
XX
XX US2003068781-A1.
XX
XX 10-APR-2003.
XX
XX 16-SEP-2002; 2002US-00245771.
XX
XX 03-AUG-1999; 99US-0146843P.
XX 15-MAY-2000; 2000MO-US013358.
XX 29-AUG-2001; 2001MO-US027099.
XX 18-JUL-2002; 2002US-00197942.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Eaton DJ, Filvaroff E, Goddard A, Grimaldi JC;
XX Gunney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
XX Fong S;
XX
XX WPI; 2003-625486/59.
XX P-PSDB; ADA06955.
XX
XX Novel secreted and transmembrane polypeptides, PRO polypeptides useful
XX for stimulating proliferation or differentiation of chondrocyte cells and
XX inducing endothelial cell tube formation.
XX
XX Claim 2; Fig 111; 307pp; English.

XX The invention relates to isolated human PRO polypeptides (secreted and

transmembrane polypeptides) and the polynucleotides encoding them. The invention also relates to an antibody which specifically binds to a PRO polypeptide, a method for stimulating the release of tumor necrosis factor-alpha (TNF-alpha) from human blood, a method for stimulating the proliferation or differentiation of chondrocyte cells and a method for detecting the presence of a tumour in a mammal (e.g. adrenal, lung, colon, breast, prostate, rectal, cervical and liver tumours). The polynucleotides are useful in molecular biology, including uses as hybridisation probes, in chromosome and gene mapping, in generating antisense RNA and DNA and in gene therapy. The polynucleotides may also be used in preparing PRO polypeptides by recombinant techniques and in generating either transgenic animals or knock-out animals which are useful in the development and screening of therapeutically useful reagents. The PRO polypeptides or antibodies are used in preparing a medicament for treating a condition responsive to the polypeptides or antibodies, such as tumours, for stimulating and inhibiting proliferation of human microvascular endothelial cells, for inducing endothelial cell tube formation and for treating various bone and/or cartilage disorders such as sports injuries and arthritis. PRO polypeptides which stimulate the release of proteoglycans from cartilage are useful for treating sports-related joint problems, articular cartilage defects, osteoarthritis and rheumatoid arthritis. PRO polypeptides are also useful for treating various mammalian haemoglobin-associated disorders such as various thalassaemias and conditions which may benefit from enhanced local immune system cell infiltration. This sequence represents a human PRO polynucleotide of the invention.

Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 GCCCGTGTGGAGAGG 43
437 GCCCGTGTGGAGAGG 455

RESULT 46
ADA08442
ID ADA08442 standard; cDNA, 943 BP.

ADA08442;
06-NOV-2003 (first entry)

Novel human secreted and transmembrane protein PRO21434 cDNA.

ss; gene; osteopathic; anti-rheumatic; antiarthritic; gene therapy;
cell proliferation stimulator;
chondrocyte cell differentiation stimulator;
secreted and transmembrane protein; PRO; human; PRO133; PRO20080;
PRO21383; human microvascular endothelial cell proliferation; PRO6071;
PRO4487; PRO6006; PRO240; PRO256; PRO1002; PRO4316; tumour;
adrenal tumour; lung tumour; colon tumour; breast tumour;
prostate tumour; rectal tumour; cervical tumour; liver tumour;
sports-related joint problem; articular cartilage defect; osteoarthritis;
rheumatoid arthritis; tissue typing.

Homo sapiens.
US2003068783-A1.

10-APR-2003.

16-SEP-2002; 2002US-00245883.

09-MAY-2001; 2001US-0290589P.
29-AUG-2001; 2001WO-US027099.
18-JUL-2002; 2002US-00197942.

(GERTH) GENENTECH INC.

Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z;
Fong S;
WPI; 2003-625486/59.
P-PSDB; ADA08443.

Novel isolated PRO133, PRO20080 or PRO21383 polypeptide useful for stimulating proliferation of human microvascular endothelial cells, and PRO6018 polypeptide useful for stimulating proliferation of chondrocyte cells.

Claim 2; Fig 111; 308pp; English.

The invention describes an isolated PRO (secreted and transmembrane) polypeptide (I). PRO6018 polypeptide is useful for stimulating the proliferation or differentiation of chondrocyte cells. PRO133, PRO20080, or PRO21383 polypeptide is useful for stimulating the proliferation of human microvascular endothelial cells. PRO6071, PRO4487, or PRO6006 polypeptide is useful for inhibiting the proliferation of human microvascular endothelial cells. PRO polypeptides such as PRO240, PRO256, PRO698, PRO1002, PRO4316, etc., are useful for detecting the presence of tumour in a mammal which involves comparing the level of expression of the above mentioned polypeptides in a test sample of cells taken from the mammal, and a control sample of normal cells of the same cell type, where a higher level of expression of the PRO polypeptide in the test sample as compared to the control sample is indicative of the presence of tumour in the mammal, the tumour being adrenal tumour, lung tumour, colon tumour, breast tumour, prostate tumour, rectal tumour, cervical tumour or liver tumour. PRO6018 polypeptide is useful for treating sports-related joint problems, articular cartilage defects, osteoarthritis, or rheumatoid arthritis. PRO polypeptides are useful as molecular weight markers for protein electrophoresis. (I) is also useful for screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). The polynucleotide (II) encoding (I) is useful as hybridisation probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid is also be useful for the preparation of PRO polypeptides. The full-length native sequence of PRO gene or its portions may be used as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs. Nucleotide sequences encoding PRO can also be used to construct hybridisation probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. (II) encoding (I) or its modified forms can also be used to generate either transgenic animals or knockout animals which in turn are useful in the development and screening of therapeutically useful reagents. (II) encoding PRO polypeptides are also useful in gene therapy techniques to treat conditions associated with aberrant expression or activity of (I). The PRO polypeptides and nucleic acid molecules are useful for tissue typing. This sequence encodes a novel human secreted and transmembrane PRO polypeptide.

Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

25 GCCCGTGTGGAGAGG 43
437 GCCCGTGTGGAGAGG 455

RESULT 47

ADB9735
ID ADB9735 standard; cDNA, 943 BP.

ADB9735;

04-DEC-2003 (first entry)

Human PRO polynucleotide SEQ ID 111.

KM Human, gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
 KM tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KM adrenal; lung; colon; breast; prostate; cervix; liver; cancer;
 KM microvascular endothelial cell; endothelial cell tube formation;
 KM sports-related joint problem; articular cartilage defect; osteoarthritis;
 KM rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
 XX Homo sapiens.
 OS US2003082726-A1.
 PN US2003082726-A1.
 XX 01-MAY-2003.
 PD 16-SEP-2002; 2002US-00245185.
 PF 07-JUL-1998; 98US-0091978P.
 PR 02-JUN-1999; 99WO-US012252.
 PR 25-AUG-1999; 99US-00380137.
 PR 29-AUG-2001; 2001WO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX (GERTH) GENENTECH INC.
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WT, Zhang Z;
 PI Fong S;
 XX WPI: 2003-743897/70.
 DR P-PSDB; ADB99736.
 PT New isolated, secreted and transmembrane PRO polypeptides and nucleic
 PT acids, useful for diagnosing, preventing and/or treating tumors, such as
 PT lung, colon, breast, prostate, rectal, kidney or liver tumors.
 XX Claim 2; Fig 111; 308pp; English.
 PS The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
 CC colon, breast, prostate, rectal, cervical and liver tumors). The
 CC polynucleotides are useful in molecular biology, including uses as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
 CC be used in preparing PRO polypeptides by recombinant techniques and in
 CC generating either transgenic animals or knock-out animals which are
 CC useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides or antibodies are used in preparing a
 CC medicament for treating a condition responsive to the polypeptides or
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation
 CC of human microvascular endothelial cells, for inducing endothelial cell
 CC tube formation and for treating sports-related joint problems, articular
 CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
 CC represents a human PRO polynucleotide of the invention.
 XX Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
 SQ
 Query Match 2.1%; Score 19; DB 9; Length 943;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 25 GCCCGGTGCTGGAGAGG 43
 Db 437 GCCCGGTGCTGGAGAGG 455
 RESULT 48
 ADB87018
 ID ADB87018 standard; CDNA, 943 BP.
 XX

AC ADB87018;
 XX 04-DEC-2003 (first entry)
 DT Human PRO polynucleotide #56.
 DE
 XX Human, gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
 KM tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; tumour;
 KM adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
 KM microvascular endothelial cell; endothelial cell tube formation;
 KM sports-related joint problem; articular cartilage defect; osteoarthritis;
 KM rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
 XX Homo sapiens.
 OS US2003082726-A1.
 PN US2003082726-A1.
 XX 01-MAY-2003.
 PD 12-SEP-2002; 2002US-00243095.
 PF 29-AUG-2001; 2001WO-US027099.
 PR 18-JUL-2002; 2002US-00197942.
 XX (GERTH) GENENTECH INC.
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WT, Zhang Z;
 PI Fong S;
 XX WPI: 2003-743897/70.
 DR P-PSDB; ADB87019.
 PT New isolated, secreted and transmembrane PRO polypeptides and nucleic
 PT acids, useful for diagnosing, preventing and/or treating tumors, such as
 PT lung, colon, breast, prostate, rectal, kidney or liver tumors.
 XX Claim 2; Fig 111; 308pp; English.
 PS The invention relates to isolated human PRO polypeptides (secreted and
 CC transmembrane polypeptides) and the polynucleotides encoding them. The
 CC invention also relates to an antibody which specifically binds to a PRO
 CC polypeptide, a method for stimulating the release of tumour necrosis
 CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
 CC proliferation or differentiation of chondrocyte cells and a method for
 CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
 CC colon, breast, prostate, rectal, cervical and liver tumors). The
 CC polynucleotides are useful in molecular biology, including uses as
 CC hybridisation probes, in chromosome and gene mapping, in generating
 CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
 CC be used in preparing PRO polypeptides by recombinant techniques and in
 CC generating either transgenic animals or knock-out animals which are
 CC useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides or antibodies are used in preparing a
 CC medicament for treating a condition responsive to the polypeptides or
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation
 CC of human microvascular endothelial cells, for inducing endothelial cell
 CC tube formation and for treating sports-related joint problems, articular
 CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
 CC represents a human PRO polynucleotide of the invention.
 XX Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
 SQ
 Query Match 2.1%; Score 19; DB 9; Length 943;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 25 GCCCGGTGCTGGAGAGG 43
 Db 437 GCCCGGTGCTGGAGAGG 455
 RESULT 49

ADB6173
ID ADB6173 standard; cDNA; 943 BP.
XX
AC ADB6173;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA encoding secreted/transmembrane polypeptide PRO21434.
XX
KW ss; gene; human; PRO; secreted protein; transmembrane protein;
KW endothelial cell tube formation; chondrocyte cell differentiation;
KW microvascular endothelial cell; tumour; lung tumour; colon tumour;
KW breast tumour; prostate tumour; rectal tumour; kidney tumour;
KW liver tumour; cytosolic; vaccine.
XX
OS Homo sapiens.
XX
PN US2003082729-A1.
XX
PD 01-MAY-2003.
XX
PF 16-SEP-2002; 2002US-00245427.
XX
PR 02-FEB-2001; 2001US-0266421P.
PR 09-FEB-2001; 2001US-0267623P.
PR 29-AUG-2001; 2001WO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
PA (GERTH) GENENTECH INC.
XX
PI Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX
DR WPI: 2003-766917/74.
DR P-PSDB; ADB6174.
XX
PT New isolated, secreted and transmembrane PRO polypeptides and nucleic
PT acid, useful for diagnosing, preventing and/or treating tumours, such as
PT lung, colon, breast, prostate, rectal, kidney or liver tumours.
XX
PS Claim 2; Fig 111; 308bp; English.
XX
CC The invention relates to an isolated secreted/transmembrane (PRO)
CC polypeptide, having at least 80% sequence identity to a sequence selected
CC from any one of the 57 amino acid sequences given in specification, or to
CC a sequence encoded by a nucleic acid molecule selected from any one of
CC the nucleic acids deposited under any of the ATCC accession numbers given
CC in specification, or a sequence having at least 80% identity to PRO
CC lacking its associated signal peptide, an extracellular domain of PRO
CC with or without its associated signal peptide. Also included are vectors,
CC transformed host cells, anti-PRO antibodies, the nucleic acids encoding
CC PRO, PRO fusion proteins, inducing endothelial cell tube formation (by
CC administering PRO281, PRO1560, PRO189, PRO4499, PRO6308, PRO6000,
CC PRO10275, PRO21207, PRO20933 or PRO34274 polypeptide or its agonist) and
CC an oligonucleotide probe derived from any one of the above nucleotide
CC sequences. PRO6018 polypeptide is useful for stimulating the
CC proliferation or differentiation of chondrocyte cells. PRO1313, PRO20080
CC and PRO21383 polypeptides are useful for stimulating the proliferation of
CC human microvascular endothelial cells. PRO6071, PRO4487 and PRO6006
CC polypeptides are useful for inhibiting the proliferation of human
CC microvascular endothelial cells. PRO polypeptides are useful for
CC detecting the presence of tumour in a mammal, including tumours of lung,
CC colon, breast, prostate, rectal, kidney and liver. PRO281, PRO1560,
CC PRO189, PRO4499, PRO6308, PRO6000, PRO10275, PRO21207, PRO20933 and
CC PRO34274 polypeptides are useful for inducing endothelial cell tube
CC formation. PRO or the antibody are useful in the preparation of a
CC medicament for treating a condition responsive to PRO polypeptide. The
CC oligonucleotide probes are useful for isolating genomic and cDNA
CC nucleotide sequences, for measuring or detecting the expression of an
CC associated gene, and as antisense probes. PRO nucleic acid is useful as a
CC hybridisation probe, in chromosome and gene mapping, in the generation of
CC antisense RNA and DNA, and for the preparation PRO polypeptides. PRO and

CC PRO nucleic acid are useful as therapeutic agents, e.g. vaccines. The
CC present sequence encodes a PRO protein.
XX
SQ Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;
XX
Query Match 2.1%; Score 19; DB 9; Length 943;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Gy 25 GCGCCGCTGGGAGAGG 43
Db 437 GCGCCGCTGGGAGAGG 455
RESULT 50
ADB99851
ID ADB99851 standard; cDNA; 943 BP.
XX
AC ADB99851;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human PRO polynucleotide SEQ ID 111.
XX
KW Human; gene; ss; PRO; secreted polypeptide; transmembrane polypeptide;
KW tumour necrosis factor-alpha; blood; chondrocyte cell; tumour;
KW adrenal; lung; colon; breast; prostate; rectum; cervix; liver; cancer;
KW microvascular endothelial cell; endothelial cell tube formation;
KW sports-related joint problem; articular cartilage defect; osteoarthritis;
KW rheumatoid arthritis; osteopathic; antirheumatic; antiarthritic.
XX
OS Homo sapiens.
XX
PN US2003073192-A1.
XX
PD 17-APR-2003.
XX
PF 16-SEP-2002; 2002US-00245055.
XX
PR 20-JUL-1999; 99US-0144758P.
PR 28-JUL-1999; 99US-0146222P.
PR 30-MAR-2000; 2000WO-US008439.
PR 02-JUN-2000; 2000WO-US015264.
PR 29-AUG-2001; 2001WO-US027099.
PR 18-JUL-2002; 2002US-00197942.
XX
PA (GERTH) GENENTECH INC.
XX
PI Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
PI Gurney AL, Smith V, Stephan JP, Watanabe CK, Wood WI, Zhang Z,
PI Fong S;
XX
DR WPI: 2003-743813/70.
DR P-PSDB; ADB99852.
XX
PT New isolated secreted and transmembrane PRO polypeptides, e.g. PRO20080
PT or PRO21383, useful in stimulating the proliferation or differentiation
PT of chondrocyte cells and detecting the presence of a tumor in a mammal.
XX
PS Claim 2; Fig 111; 308bp; English.
XX
CC The invention relates to isolated human PRO polypeptides (secreted and
CC transmembrane polypeptides) and the polynucleotides encoding them. The
CC invention also relates to an antibody which specifically binds to a PRO
CC polypeptide, a method for stimulating the release of tumour necrosis
CC factor-alpha (TNF-alpha) from human blood, a method for stimulating the
CC proliferation or differentiation of chondrocyte cells and a method for
CC detecting the presence of a tumour in a mammal (e.g. adrenal, lung,
CC colon, breast, prostate, rectal, cervical and liver tumours). The
CC polynucleotides are useful in molecular biology, including uses as
CC hybridisation probes, in chromosome and gene mapping, in generating
CC antisense RNA and DNA and in gene therapy. The polynucleotides may also
CC be used in preparing PRO polypeptides by recombinant techniques and in

CC generating either transgenic animals or knock-out animals which are
 CC useful in the development and screening of therapeutically useful
 CC reagents. The PRO polypeptides or antibodies are used in preparing a
 CC medicament for treating a condition responsive to the polypeptides or
 CC antibodies, such as tumours, for stimulating and inhibiting proliferation
 CC of human microvascular endothelial cells, for inducing endothelial cell
 CC tube formation and for treating sports-related joint problems, articular
 CC cartilage defects, osteoarthritis and rheumatoid arthritis. This sequence
 CC represents a human PRO polynucleotide of the invention.

XX
 SQ Sequence 943 BP; 152 A; 306 C; 299 G; 186 T; 0 U; 0 Other;

Query Match 2.1%; Score 19; DB 10; Length 943;

Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GCGCCGCTGGGAGGAGG 43

Db 437 GCGCCGCTGGGAGGAGG 455

Search completed: March 23, 2005, 20:49:52
 Job time : 534.777 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 24, 2005, 00:19:51 ; Search time 817.188 Seconds

(without alignments)
6628.866 Million cell updates/sec

Title: US-10-030-271-1

Perfect score: 909

Sequence: 1 atggcgctatccgggtcgcac.....tcagctgtgatgagctgcac 909

Scoring table: OLIGO_NUC

Searched: 5552208 seqs, 297965951 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCRU5_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	858	94.4	1230	16	US-10-296-539-4
2	858	94.4	1924	14	US-10-001-254-17
3	858	94.4	2045	14	US-10-013-477-2
4	858	94.4	2045	15	US-10-106-698-349
5	768	84.5	2044	9	US-09-925-302-315
6	768	84.5	2044	10	US-09-925-302-315
7	538	59.2	1966	9	US-09-822-830A-390
8	531	58.4	1570	17	US-10-296-115-481
9	321	35.3	1554	9	US-09-799-777-103
10	281	30.9	1067	15	US-10-037-270-853
11	281	30.9	1067	17	US-10-117-722-853

12	277	30.5	111084	18	US-10-723-860-1627	Sequence 1627, Ap
13	252	27.7	303	14	US-10-001-254-7	Sequence 7, Appl
14	151	16.6	451	10	US-09-918-995-28421	Sequence 28421, A
15	96	10.6	484	10	US-09-918-995-12268	Sequence 12268, A
16	22	2.4	838	10	US-09-814-353-18768	Sequence 18768, A
17	21	2.3	1439	17	US-09-814-353-13540	Sequence 13540, A
18	21	2.3	1439	17	US-10-425-114-837	Sequence 837, App
19	21	2.3	1496	18	US-10-425-115-14144	Sequence 14144, A
20	20	2.2	411	18	US-10-425-115-46300	Sequence 46300, A
21	20	2.2	424	18	US-10-437-963-90512	Sequence 90512, A
22	20	2.2	2304	15	US-10-156-761-6280	Sequence 6280, Ap
23	20	2.2	2610	16	US-10-246-330-1	Sequence 1, Appl
24	20	2.2	9025608	15	US-10-156-761-1	Sequence 1, Appl
25	19	2.1	201	18	US-10-719-993-13899	Sequence 13899, A
26	19	2.1	625	13	US-10-027-632-185402	Sequence 185402, A
27	19	2.1	625	13	US-10-027-632-185403	Sequence 185403, A
28	19	2.1	625	17	US-10-027-632-185402	Sequence 185402, A
29	19	2.1	625	17	US-10-027-632-185403	Sequence 185403, A
30	19	2.1	682	13	US-10-027-632-209406	Sequence 209406, A
31	19	2.1	682	13	US-10-027-632-209407	Sequence 209407, A
32	19	2.1	682	17	US-10-027-632-209406	Sequence 209406, A
33	19	2.1	682	17	US-10-027-632-209407	Sequence 209407, A
34	19	2.1	943	14	US-10-245-752-111	Sequence 111, App
35	19	2.1	943	14	US-10-245-859-111	Sequence 111, App
36	19	2.1	943	14	US-10-245-103-111	Sequence 111, App
37	19	2.1	943	14	US-10-245-107-111	Sequence 111, App
38	19	2.1	943	14	US-10-245-143-111	Sequence 111, App
39	19	2.1	943	14	US-10-245-771-111	Sequence 111, App
40	19	2.1	943	14	US-10-245-851-111	Sequence 111, App
41	19	2.1	943	14	US-10-245-883-111	Sequence 111, App
42	19	2.1	943	14	US-10-237-535-111	Sequence 111, App
43	19	2.1	943	14	US-10-238-183-111	Sequence 111, App
44	19	2.1	943	14	US-10-238-283-111	Sequence 111, App
45	19	2.1	943	14	US-10-238-370-111	Sequence 111, App
46	19	2.1	943	14	US-10-245-055-111	Sequence 111, App
47	19	2.1	943	14	US-10-245-147-111	Sequence 111, App
48	19	2.1	943	14	US-10-245-730-111	Sequence 111, App
49	19	2.1	943	14	US-10-245-739-111	Sequence 111, App
50	19	2.1	943	14	US-10-246-210-111	Sequence 111, App
51	19	2.1	943	14	US-10-239-196-111	Sequence 111, App
52	19	2.1	943	14	US-10-243-024-111	Sequence 111, App
53	19	2.1	943	14	US-10-243-409-111	Sequence 111, App
54	19	2.1	943	14	US-10-243-621-111	Sequence 111, App
55	19	2.1	943	14	US-10-245-880-111	Sequence 111, App
56	19	2.1	943	14	US-10-245-033-111	Sequence 111, App
57	19	2.1	943	14	US-10-243-099-111	Sequence 111, App
58	19	2.1	943	14	US-10-243-185-111	Sequence 111, App
59	19	2.1	943	14	US-10-245-427-111	Sequence 111, App
60	19	2.1	943	14	US-10-245-427-111	Sequence 111, App
61	19	2.1	943	14	US-10-245-770-111	Sequence 111, App
62	19	2.1	943	14	US-10-245-877-111	Sequence 111, App
63	19	2.1	943	14	US-10-246-976-111	Sequence 111, App
64	19	2.1	943	14	US-10-243-320-111	Sequence 111, App
65	19	2.1	943	14	US-10-243-743-111	Sequence 111, App
66	19	2.1	943	14	US-10-242-845-111	Sequence 111, App
67	19	2.1	943	15	US-10-237-636-111	Sequence 111, App
68	19	2.1	943	15	US-10-238-325-111	Sequence 111, App
69	19	2.1	943	15	US-10-238-346-111	Sequence 111, App
70	19	2.1	943	15	US-10-238-411-111	Sequence 111, App
71	19	2.1	943	15	US-10-243-167-111	Sequence 111, App
72	19	2.1	943	15	US-10-243-386-111	Sequence 111, App
73	19	2.1	943	15	US-10-243-446-111	Sequence 111, App
74	19	2.1	943	15	US-10-245-874-111	Sequence 111, App
75	19	2.1	943	15	US-10-242-653-111	Sequence 111, App
76	19	2.1	943	15	US-10-243-167-111	Sequence 111, App
77	19	2.1	943	15	US-10-243-386-111	Sequence 111, App
78	19	2.1	943	15	US-10-244-947-111	Sequence 111, App
79	19	2.1	943	15	US-10-244-966-111	Sequence 111, App
80	19	2.1	943	15	US-10-244-990-111	Sequence 111, App
81	19	2.1	943	15	US-10-245-079-111	Sequence 111, App
82	19	2.1	943	15	US-10-245-127-111	Sequence 111, App
83	19	2.1	943	15	US-10-245-207-111	Sequence 111, App
84	19	2.1	943	15	US-10-245-646-111	Sequence 111, App

```
85 19 2.1 943 15 US-10-245-695-111 Sequence 111, App
86 19 2.1 943 15 US-10-245-699-111 Sequence 111, App
87 19 2.1 943 15 US-10-245-737-111 Sequence 111, App
88 19 2.1 943 15 US-10-245-878-111 Sequence 111, App
89 19 2.1 943 15 US-10-245-890-111 Sequence 111, App
90 19 2.1 943 15 US-10-245-899-111 Sequence 111, App
91 19 2.1 943 15 US-10-245-900-111 Sequence 111, App
92 19 2.1 943 15 US-10-247-058-111 Sequence 111, App
93 19 2.1 943 15 US-10-245-454-111 Sequence 111, App
94 19 2.1 943 15 US-10-237-471-111 Sequence 111, App
95 19 2.1 943 15 US-10-238-261-111 Sequence 111, App
96 19 2.1 943 15 US-10-238-324-111 Sequence 111, App
97 19 2.1 943 15 US-10-241-860-111 Sequence 111, App
98 19 2.1 943 15 US-10-242-172-111 Sequence 111, App
99 19 2.1 943 15 US-10-242-652-111 Sequence 111, App
100 19 2.1 943 15 US-10-242-990-111 Sequence 111, App
```

ALIGNMENTS

RESULT 1

```
US-10-296-539-4
; Sequence 4, Application US/10296539
; Publication No. US20030165933A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: REGULATORS OF APOPTOSIS
; FILE REFERENCE: PI-0307 PCT
; CURRENT APPLICATION NUMBER: US/10/296,539
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 60/209,407; 60/250,326
; PRIOR FILING DATE: 2000-06-01; 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030165933A1 3102521CB1
US-10-296-539-4
```

```
Query Match 94.4%; Score 858; DB 16; Length 1230;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGTATCCGGGTGACCCCGGCCCGTGTGAGGAGATGATGCTCTGACTAC 60
Db 127 ATGGCGTATCCGGGTGACCCCGGCCCGTGTGAGGAGATGATGCTCTGACTAC 186
QY 61 TACGGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGCGCAACTGACCGATGC 120
Db 187 TACGGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGCGCAACTGACCGATGC 246
QY 121 GAACTGAGAGTCTCTGGGCTTTTGTGCTGATGAGGCTCTCTGGCGCCGCGGAGCTTATGCC 180
Db 247 GAACTGAGAGTCTCTGGGCTTTTGTGCTGATGAGGCTCTCTGGCGCCGCGGAGCTTATGCC 306
QY 181 CGGGCCGCGAGCGGCTTGAAGCTCTGCTGAGCTGAGGCGCGCGGCGAGTGCGGCGAG 240
Db 307 CGGGCCGCGAGCGGCTTGAAGCTCTGCTGAGCTGAGGCGCGCGGCGAGTGCGGCGAG 366
QY 241 AGCAACTGCGGCTGTGCTGGGCGAACTCTGCGGTGTGCGCGCGCGCAAGCACTGTGCGG 300
```

```
|||||
Db 367 AGCAACTGCGGCTGTGCTGGGCGAACTCTGCGGTGTGCGCGCGCGCAAGCACTGTGCGG 426
QY 301 CACCTGCGCGAGAGCGGCGCGGCGGCGAGTGTCTCCAGAAAGCTATAGTATGGCACCTGC 360
Db 427 CACCTGCGCGAGAGCGGCGCGGCGGCGAGTGTCTCCAGAAAGCTATAGTATGGCACCTGC 486
QY 361 AGCTCTTCAAGAGAGCAGAGGGGTAGCTGCGCTGCGGCGAGTCAAGAGTTTGTGA 420
Db 487 AGCTCTTCAAGAGAGCAGAGGGGTAGCTGCGCTGCGGCGAGTCAAGAGTTTGTGA 546
QY 421 AATTCTCAGAGGAGTCAAGTGGAGAGAGAGCTCCCTCCCAACCAAGGCGAGCGGAGT 480
Db 547 AATTCTCAGAGGAGTCAAGTGGAGAGAGAGCTCCCTCCCAACCAAGGCGAGCGGAGT 606
QY 481 CGGGCGCGGCGCGAGGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCCGACCGCACCCAG 540
Db 607 CGGGCGCGGCGCGAGGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCCGACCGCACCCAG 666
QY 541 CAGCACTCAGAGCGCGCGAGACCTTCTCTGAAGGCAAGTGAAGCTGTGACATCCGGCTC 600
Db 667 CAGCACTCAGAGCGCGCGAGACCTTCTCTGAAGGCAAGTGAAGCTGTGACATCCGGCTC 726
QY 601 CGGGTTGAGCAGAGTACTGTGCGAGCATGGGCCAGCCTTTGAGAGCGGCGTGGCATCCCG 660
Db 727 CGGGTTGAGCAGAGTACTGTGCGAGCATGGGCCAGCCTTTGAGAGCGGCGTGGCATCCCG 786
QY 661 CGGGCCCGAGGCGGTGCGGCGGCGAGCTGAGACGTGTTTGGCGAGGCGCACCGAGTGGCGG 720
Db 787 CGGGCCCGAGGCGGTGCGGCGGCGAGCTGAGACGTGTTTGGCGAGGCGCACCGAGTGGCGG 846
QY 721 TCAGAGGAGCTGTGGCTCTGTGTGTTTGTGACATCAAGTTTCAAGAGTCTCTATCTGAG 780
Db 847 TCAGAGGAGCTGTGGCTCTGTGTGTTTGTGACATCAAGTTTCAAGAGTCTCTATCTGAG 906
QY 781 GCTTTTGGGCGACATCACTGAGTGGCGGCTGTGCGAGGCTGTGGGGCGGTGTTCTTG 840
Db 907 GCTTTTGGGCGACATCACTGAGTGGCGGCTGTGCGAGGCTGTGGGGCGGTGTTCTTG 966
QY 841 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGGAGGCTGTTGCGCTGTGAGTGTGAGT 900
Db 967 ACTGAGGCGCTGCGAGAGGCTGTGGGCGGGAGGCTGTTGCGCTGTGAGTGTGAGT 1026
QY 901 GAGGCTGAC 909
Db 1027 GAGGCTGAC 1035
```

RESULT 2

```
US-10-001-254-17
; Sequence 17, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Reed, John C.
; APPLICANT: Godzicki, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Steiner-Bliwien, Frank
; TITLE OF INVENTION: No. US20030049702A1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 1924
; TYPE: DNA
```


ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (91)...(1044)
NAME/KEY: misc_feature
LOCATION: (1)...(1900)
OTHER INFORMATION: n = A,T,C or G
US-10-001-254-17

Query Match 94.4%; Score 858; DB 14; Length 1924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGCGCTATCCGGGTGAGCCCGCCCGCTGCTGGAGAGAGATGAGTCTGAGACTAC 60
DB 91 ATGGCGCTATCCGGGTGAGCCCGCCCGCTGCTGGAGAGAGATGAGTCTGAGACTAC 150
QY 61 TACGGAGATGCTGTGCTTCAACCGTATGTTCAAGGTGGTGGCGGAGCACTGACCGAGTGC 120
DB 151 TACGGAGATGCTGTGCTTCAACCGTATGTTCAAGGTGGTGGCGGAGCACTGACCGAGTGC 210
QY 121 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCCGCGAGGCTTAAGCC 180
DB 211 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCCGCGAGGCTTAAGCC 270
QY 181 CGGGCCCGGAGCGGCTAGAGCTCTGCTGAGAGCTGAGAGCGCGGCGGCACTGCGGAG 240
DB 271 CGGGCCCGGAGCGGCTAGAGCTCTGCTGAGAGCTGAGAGCGCGGCGGCACTGCGGAG 330
QY 241 AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCGTGGCTGAGCGCGCAAGCTTGTGCG 300
DB 331 AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCGTGGCTGAGCGCGCAAGCTTGTGCG 390
QY 301 CACCTGCGCGGCAAGCGGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGAGCACTCC 360
DB 391 CACCTGCGCGGCAAGCGGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGAGCACTCC 450
QY 361 AGCTCTTCAAGAGAGAGAGAGAGTGTGCTGCGCGTGGCTGAGAGAGTGTGCA 420
DB 451 AGCTCTTCAAGAGAGAGAGAGAGTGTGCTGCGCGTGGCTGAGAGAGTGTGCA 510
QY 421 AATTCTCAGACAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
DB 511 AATTCTCAGACAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 570
QY 481 CGGGCCCGGCGCAAGTGTGTCAGACAGCGCGGAGAGAGGAGCCCAAGCGGCACTCC 540
DB 571 CGGGCCCGGCGCAAGTGTGTCAGACAGCGCGGAGAGAGGAGCCCAAGCGGCACTCC 630
QY 541 CAGAGATCAGAGCGGCGGCAAGCTTCTCTGAGAGCAAGTGTGATCTGAGTCTC 600
DB 631 CAGAGATCAGAGCGGCGGCAAGCTTCTCTGAGAGCAAGTGTGATCTGAGTCTC 690
QY 601 CGGGTTGAGAGAGAGTCTGAGAGATGAGGCGAGCTTGGAGAGAGGAGTGTGCACTCC 660
DB 691 CGGGTTGAGAGAGAGTCTGAGAGATGAGGCGAGCTTGGAGAGAGGAGTGTGCACTCC 750
QY 661 CGGGCCCGGCGCTGAGCGCGGCAAGCTGTGTTGGAGAGGCGCAAGCTGAGTCTG 720
DB 751 CGGGCCCGGCGCTGAGCGCGGCAAGCTGTGTTGGAGAGGCGCAAGCTGAGTCTG 810
QY 721 TCAAGGAGAGCTGAGGCTCTGTGTTGGATCAAGTGTGATCAAGGCTCTCTATCTG 780
DB 811 TCAAGGAGAGCTGAGGCTCTGTGTTGGATCAAGTGTGATCAAGGCTCTCTATCTG 870
QY 781 GCCTTCTGAGGCGAGCTACCTGAGTGGCGCTGCTGAGAGGCGCTGCGGAGGCTTCTG 840
DB 871 GCCTTCTGAGGCGAGCTACCTGAGTGGCGCTGCTGAGAGGCGCTGCGGAGGCTTCTG 930
QY 841 ACTGAGGCGCTGAGAGAGGCTGTGGGCGGAGAGGCTGTTGCGCTGCTGAGTGTGAT 900
DB 931 ACTGAGGCGCTGAGAGAGGCTGTGGGCGGAGAGGCTGTTGCGCTGCTGAGTGTGAT 990
```

QY 901 GAGGCTGAC 909
DB 991 GAGGCTGAC 999

RESULT 3

```
US-10-013-477-2
; Sequence 2, Application US/10013477
; Publication No. US2003049732A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: Apoptosis Related Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT002P1
; CURRENT APPLICATION NUMBER: US/10/013,477
; PRIORITY FILING DATE: 2001-12-13
; PRIORITY APPLICATION NUMBER: 09/669,445
; PRIORITY FILING DATE: 2000-09-25
; PRIORITY APPLICATION NUMBER: PCT/US00/06642
; PRIORITY FILING DATE: 2000-03-15
; PRIORITY APPLICATION NUMBER: 60/126,018
; PRIORITY FILING DATE: 1999-03-24
; PRIORITY APPLICATION NUMBER: 60/139,638
; PRIORITY FILING DATE: 1999-06-17
; PRIORITY APPLICATION NUMBER: 60/149,449
; PRIORITY FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-013-477-2
```

Query Match 94.4%; Score 858; DB 14; Length 2045;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGCGCTATCCGGGTGAGCCCGCCCGCTGCTGGAGAGAGATGAGTCTGAGACTAC 60
DB 121 ATGGCGCTATCCGGGTGAGCCCGCCCGCTGCTGGAGAGAGATGAGTCTGAGACTAC 180
QY 61 TACGGAGATGCTGTGCTTCAACCGTATGTTCAAGGTGGTGGCGGAGCACTGACCGAGTGC 120
DB 181 TACGGAGATGCTGTGCTTCAACCGTATGTTCAAGGTGGTGGCGGAGCACTGACCGAGTGC 240
QY 121 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCCGCGAGGCTTAAGCC 180
DB 241 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCGCCGCGAGGCTTAAGCC 300
QY 181 CGGGCCCGGAGCGGCTAGAGCTCTGCTGAGAGCTGAGAGCGCGGCGGCACTGCGGAG 240
DB 301 CGGGCCCGGAGCGGCTAGAGCTCTGCTGAGAGCTGAGAGCGCGGCGGCACTGCGGAG 360
QY 241 AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCGTGGCTGAGCGCGCAAGCTTGTGCG 300
DB 361 AGCAACCTGCGGCTGCTGAGGCAACTCTGCGCGTGGCTGAGCGCGCAAGCTTGTGCG 420
QY 301 CACCTGCGCGGCAAGCGGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGAGCACTCC 360
DB 421 CACCTGCGCGGCAAGCGGCGCGGCAAGTGTCTCAGAACGCTATAGCTATGAGCACTCC 480
QY 421 GAGCTCTTCAAGAGAGAGAGAGTGTGCTGCGCGTGGCTGAGAGCTGAGAGCTTCTGCA 420
DB 481 GAGCTCTTCAAGAGAGAGAGAGTGTGCTGCGCGTGGCTGAGAGCTTCTGCA 540
QY 481 AATTCTCAGACAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 480
DB 541 AATTCTCAGACAGGCTCAGTGGAGAGACAGGCTCCCCCAACCAAGCGGAGCGGAGT 600
QY 481 CGGGCCCGGCGCAAGTGTGTCAGACAGCGCGGAGAGAGGAGCCCAAGCTGCACTCC 540
DB 601 CGGGCCCGGCGCAAGTGTGTCAGACAGCGCGGAGAGAGGAGCCCAAGCTGCACTCC 660
```

```
QY 541 CAGAGTCAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGAAGCTGTGAATCCGGCTC 600
|
|
|
Db 661 CAGAGTCAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGAAGCTGTGAATCCGGCTC 720
|
|
|
QY 601 CGGGTTTCAGACAGATCTCTGAGAGCATAGGGCCAGCTTGGAGCAGGGCCGTGGCATCCCG 660
|
|
|
Db 721 CGGGTTTCAGACAGATCTCTGAGAGCATAGGGCCAGCTTGGAGCAGGGCCGTGGCATCCCG 780
|
|
|
QY 661 CGGGCCCGCAGCGCTGGCGCGGAGCTGAGAGTGTGTTGGAGGCAACCGAGTGTCTCG 720
|
|
|
Db 781 CGGGCCCGCAGCGCTGGCGCGGAGCTGAGAGTGTGTTGGAGGCAACCGAGTGTCTCG 840
|
|
|
QY 721 TCAAGGAGCTGGGCTGTGTGTGTGTGATCAAGTTCAGAGCTCTCTAATCTGGAC 780
|
|
|
Db 841 TCAAGGAGCTGGGCTGTGTGTGTGTGATCAAGTTCAGAGCTCTCTAATCTGGAC 900
|
|
|
QY 781 GCCTTCTGGGGGCGCATCTGAGAGTGGCGCCCTGTGAGAGCCCTGGCGGGCGTGTCTG 840
|
|
|
Db 901 GCCTTCTGGGGGCGCATCTGAGAGTGGCGCCCTGTGAGAGCCCTGGCGGGCGTGTCTG 960
|
|
|
QY 841 ACTGAGGCTCTGCGAGAGGCTGTGGGCGGGAGGCTGTTCGCTGTGATCAAGTGTGAT 900
|
|
|
Db 961 ACTGAGGCTCTGCGAGAGGCTGTGGGCGGGAGGCTGTTCGCTGTGATCAAGTGTGAT 1020
|
|
|
QY 901 GAGGCTGAC 909
|
|
|
Db 1021 GAGGCTGAC 1029
|
|
|
```

RESULT 4

```
US-10-106-698-349
; Sequence 349, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 349
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-349
```

Query Match 94.4%; Score 858; DB 15; Length 2045;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 908; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGCGTATCCGGGTGACCCCGCCCGGCTGGAGGAGATGATGCTCGACTAC 60
|
|
|
Db 121 ATGAGCGTATCCGGGTGACCCCGCCCGGCTGGAGGAGATGATGCTCGACTAC 180
|
|
|
QY 61 TACGGAGTGTGTGCTTCAACCGTATGTTGAGGTGTGGCGGCACTGACGATGC 120
|
|
|
Db 181 TACGGAGTGTGTGCTTCAACCGTATGTTGAGGTGTGGCGGCACTGACGATGC 240
|
|
|
QY 121 GAGCTGAGAGCTCTGCGCTTTCGTGCTGATGAGGCTCTTGCGCGCCGAGAGCTTATGCC 180
|
|
|
Db 241 GAGCTGAGAGCTCTGCGCTTTCGTGCTGATGAGGCTCTTGCGCGCCGAGAGCTTATGCC 300
|
|
|
QY 181 CGGAGCCGCGAGCGGCTAGAGCTCTGCTGAGAGTGGAGCGCGGCGAGTGTGCGAG 240
|
|
|
Db 301 CGGAGCCGCGAGCGGCTAGAGCTCTGCTGAGAGTGGAGCGCGGCGAGTGTGCGAG 360
|
|
|
```

```
QY 241 AGCAACCTGGCGCTGTGGGCAACTCTGCGGTGTGGCCCGCAGCACTGTGCGG 300
|
|
|
Db 361 AGCAACCTGGCGCTGTGGGCAACTCTGCGGTGTGGCCCGCAGCACTGTGCGG 420
|
|
|
QY 301 CACCTGGCGCGAGAGGCGCGCCGAGTGTCTCCAGAAAGCTATAGTATGCACTCC 360
|
|
|
Db 421 CACCTGGCGCGAGAGGCGCGCCGAGTGTCTCCAGAAAGCTATAGTATGCACTCC 480
|
|
|
QY 361 AGCTTTCAAAAGAGCAGAGGAGTACTCCGCTCGCGCTCGGAGTCAAGAGTTCGA 420
|
|
|
Db 481 AGCTTTCAAAAGAGCAGAGGAGTACTCCGCTCGCGCTCGGAGTCAAGAGTTCGA 540
|
|
|
QY 421 AATTCTCAGAGGAGTCAAGTGGAGAGCAGGCTCCCCCAACCAAGCGAGCGGAGT 480
|
|
|
Db 541 AATTCTCAGAGGAGTCAAGTGGAGAGCAGGCTCCCCCAACCAAGCGAGCGGAGT 600
|
|
|
QY 481 CGGGGCGGCGCCAGTGTGTGTGTGCAAGCGCGCGGAGAGGGGCCCCAGCCGAC 540
|
|
|
Db 601 CGGGGCGGCGCCAGTGTGTGTGTGCAAGCGCGCGGAGAGGGGCCCCAGCCGAC 660
|
|
|
QY 541 CAGCAATCAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGAAGCTGTGATCCGGCTC 600
|
|
|
Db 661 CAGCAATCAGAGCCCGCAGACCTTCTCTGAAGGCAAGTGAAGCTGTGATCCGGCTC 720
|
|
|
QY 601 CGGGTTTCAGACAGATCTCTGAGAGTGGGCGCAGCTTGGAGGCGCGTGTGATCCG 660
|
|
|
Db 721 CGGGTTTCAGACAGATCTCTGAGAGTGGGCGCAGCTTGGAGGCGCGTGTGATCCG 780
|
|
|
QY 661 CGGGCCCGCAGCGCTGGCGCGGAGCTGAGAGTGTGTTGGAGGCAACCGAGTGTCTG 720
|
|
|
Db 781 CGGGCCCGCAGCGCTGGCGCGGAGCTGAGAGTGTGTTGGAGGCAACCGAGTGTCTG 840
|
|
|
QY 721 TCAAGGAGCTGGGCTGTGTGTGTGTGATCAAGTTCAGAGCTCTCTAATCTGGAC 780
|
|
|
Db 841 TCAAGGAGCTGGGCTGTGTGTGTGTGATCAAGTTCAGAGCTCTCTAATCTGGAC 900
|
|
|
QY 781 GCCTTCTGGGGGCGCATCTGAGAGTGGCGCCCTGTGAGAGCCCTGGCGGGCGTGTCTG 840
|
|
|
Db 901 GCCTTCTGGGGGCGCATCTGAGAGTGGCGCCCTGTGAGAGCCCTGGCGGGCGTGTCTG 960
|
|
|
QY 841 ACTGAGGCTCTGCGAGAGGCTGTGGGCGGGAGGCTGTTCGCTGTGATCAAGTGTGAT 900
|
|
|
Db 961 ACTGAGGCTCTGCGAGAGGCTGTGGGCGGGAGGCTGTTCGCTGTGATCAAGTGTGAT 1020
|
|
|
QY 901 GAGGCTGAC 909
|
|
|
Db 1021 GAGGCTGAC 1029
|
|
|
```

RESULT 5

```
US-09-925-302-315
; Sequence 315, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 315
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-315
```

Query Match 84.5%; Score 768; DB 9; Length 2044;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 818; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTGACCCCGGCCCGTCTGGAGAGATGATGCTTGGACTAC 60
DB 121 ATGGCGCTATCCGGGTGACCCCGGCCCGTCTGGAGAGATGATGCTTGGACTAC 180
QY 61 TACGGAGTGTCTGCTTCACTCCGATGTGAGGTGGTGGGCGGAGCACTGACGAGTGC 120
DB 181 TACGGAGTGTCTGCTTCACTCCGATGTGAGGTGGTGGGCGGAGCACTGACGAGTGC 240
QY 121 GAGCTGAGAGCTCTGCGCTTCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGCTTAC 180
DB 241 GAGCTGAGAGCTCTGCGCTTCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGCTTAC 300
QY 181 CGGGCCCGGAGCGGCTTGAAGCTCTCTGCTGAGCTGAGCGCGCGGAGCTTAC 240
DB 301 CGGGCCCGGAGCGGCTTGAAGCTCTCTGCTGAGCTGAGCGCGCGGAGCTTAC 360
QY 241 AGCAACCTGCGGCTGCTGGGGCACTCTGCGGCTGAGCGCGCGGAGCTTAC 300
DB 361 AGCAACCTGCGGCTGCTGGGGCACTCTGCGGCTGAGCGCGCGGAGCTTAC 420
QY 301 CACCTGCGGAGCGGCTGAGCGCGGAGCTCTTCAAGAGCTTATGAGCACTTCC 360
DB 421 CACCTGCGGAGCGGCTGAGCGCGGAGCTCTTCAAGAGCTTATGAGCACTTCC 480
QY 361 AGCTCTTCAAAAG 420
DB 481 AGCTCTTCAAAAG 540
QY 421 AATTCTGAG 480
DB 541 AATTCTGAG 600
QY 481 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 601 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 541 CAGAGAGTCAAG 600
DB 661 CAGAGAGTCAAG 720
QY 601 CGGGTTCGAG 660
DB 721 CGGGTTCGAG 780
QY 661 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 781 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 721 TCAAGGAGAGCTGAG 780
DB 841 TCAAGGAGAGCTGAG 900
QY 781 GCCTTCGAG 819
DB 901 GCCTTCGAG 939

RESULT 6
US-09-925-302-315
; Sequence 315, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 315
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-315

Query Match 84.5%; Score 768; DB 10; Length 2044;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 818; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTATCCGGGTGACCCCGGCCCGTCTGGAGAGATGATGCTTGGACTAC 60
DB 121 ATGGCGCTATCCGGGTGACCCCGGCCCGTCTGGAGAGATGATGCTTGGACTAC 180
QY 61 TACGGAGTGTCTGCTTCACTCCGATGTGAGGTGGTGGGCGGAGCACTGACGAGTGC 120
DB 181 TACGGAGTGTCTGCTTCACTCCGATGTGAGGTGGTGGGCGGAGCACTGACGAGTGC 240
QY 121 GAGCTGAGAGCTCTGCGCTTCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGCTTAC 180
DB 241 GAGCTGAGAGCTCTGCGCTTCTTCTGCTGATGAGGCTCTTGGCGCGCGGAGCTTAC 300
QY 181 CGGGCCCGGAGCGGCTTGAAGCTCTCTGCTGAGCTGAGCGCGCGGAGCTTAC 240
DB 301 CGGGCCCGGAGCGGCTTGAAGCTCTCTGCTGAGCTGAGCGCGCGGAGCTTAC 360
QY 241 AGCAACCTGCGGCTGCTGGGGCACTCTGCGGCTGAGCGCGCGGAGCTTAC 300
DB 361 AGCAACCTGCGGCTGCTGGGGCACTCTGCGGCTGAGCGCGCGGAGCTTAC 420
QY 301 CACCTGCGGAGCGGCTGAGCGCGGAGCTCTTCAAGAGCTTATGAGCACTTCC 360
DB 421 CACCTGCGGAGCGGCTGAGCGCGGAGCTCTTCAAGAGCTTATGAGCACTTCC 480
QY 361 AGCTCTTCAAAAG 420
DB 481 AGCTCTTCAAAAG 540
QY 421 AATTCTGAG 480
DB 541 AATTCTGAG 600
QY 481 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 601 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 541 CAGAGAGTCAAG 600
DB 661 CAGAGAGTCAAG 720
QY 601 CGGGTTCGAG 660
DB 721 CGGGTTCGAG 780
QY 661 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 781 CGGGCCCGGAGCGGCTGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 721 TCAAGGAGAGCTGAG 780
DB 841 TCAAGGAGAGCTGAG 900
QY 781 GCCTTCGAG 819
DB 901 GCCTTCGAG 939

RESULT 7
US-09-822-830A-390
; Sequence 390, Application US/09822830A
; Patent No. US20020142952A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Genetics Institute, Inc.
/ APPLICANT: Wong, Gordon G.
/ APPLICANT: Clark, Hilary
/ APPLICANT: Fechtel, Kim
/ APPLICANT: Agostino, Michael J.
/ APPLICANT: Howes, Steven H.
/ APPLICANT: Resnick, Richard J.
/ APPLICANT: Gulukota, Kamalakhar
/ APPLICANT: Graham, James R.
/ TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
/ FILE REFERENCE: GIN 6402
/ CURRENT APPLICATION NUMBER: US/09/822,830A
/ CURRENT FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/195,604
/ PRIOR FILING DATE: 2000-04-06
/ NUMBER OF SEQ ID NOS: 631
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 390
/ LENGTH: 1966
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-822-830A-390
```

```
Query Match          59.2%; Score 538; DB 9; Length 1966;
Best Local Similarity 99.8%; Pred. No. 1.3e-253;
Matches 588; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGCTATCCGGGTGACCCCGGCCCGCTGCTGGAGAGAGATGAGTGCCTGGAATAC 60
DB 50 ATGGCGCTATCCGGGTGACCCCGGCCCGCTGCTGGAGAGAGATGAGTGCCTGGAATAC 109
QY 61 TACGGAGTGTGCTGCTTCAACCGTATGTTGAGGTGGTGGGCGGCAATGACCGAGTGC 120
DB 110 TACGGAGTGTGCTGCTTCAACCGTATGTTGAGGTGGTGGGCGGCAATGACCGAGTGC 169
QY 121 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTATGCC 180
DB 170 GAGCTGAGAGCTCTGAGCTTTCTGCTGATGAGGCTCTGCGCGCGCGGAGGCTTATGCC 229
QY 181 CGGGCCCGGAGCGGCTTGAAGCTCTGCTGAGAGTGGAGCGGCGGCGAGTGGAGAG 240
DB 230 CGGGCCCGGAGCGGCTTGAAGCTCTGCTGAGAGTGGAGCGGCGGCGAGTGGAGAG 289
QY 241 AGCAAGCTGCGGCTGCTGAGGCAACTCTGCGCGTGGCGCGCGCAAGCTGCTGCGG 300
DB 290 AGCAAGCTGCGGCTGCTGAGGCAACTCTGCGCGTGGCGCGCGCAAGCTGCTGCGG 349
QY 301 CACCTGCGCGCAAGCGGCGCGCGCAAGTCTTTCAGAACGCTATAGCTATGCGACTTCC 360
DB 350 CACCTGCGCGCAAGCGGCGCGCGCAAGTCTTTCAGAACGCTATAGCTATGCGACTTCC 409
QY 361 AGCTCTTCAAGAGAGAGAGAGTGAAGTGCCTGCTGCGCGCTGCGAGTCAAGCACTTTCGA 420
DB 410 AGCTCTTCAAGAGAGAGAGAGTGAAGTGCCTGCTGCGCGCTGCGAGTCAAGCACTTTCGA 469
QY 421 AATTCTGAGCGGCTGCTGAGAGAGAGGCTCCCCCGCAACCAAGCGGCGGAGT 480
DB 470 AATTCTGAGCGGCTGCTGAGAGAGAGGCTCCCCCGCAACCAAGCGGCGGAGT 529
QY 481 CGGGCGCGGCGCAAGTGTGTGTCAGAGCGGCGCGAGAGAGAGGCGCCAGCGCAGCCAG 540
DB 530 CGGGCGCGGCGCAAGTGTGTGTCAGAGCGGCGCGAGAGAGAGGCGCCAGCGCAGCCAG 589
QY 541 CAGAGAGCAGAGCGCGCGCAAGCTTCTCTGAGAGGCAAGTGAAGCTGTG 589
DB 590 CAGAGAGCAGAGCGCGCGCAAGCTTCTCTGAGAGGCAAGTGAAGCTGTG 638
```

```
RESULT 8
US-10-296-115-481
/ Sequence 481, Application US/10296115
/ Publication No. US20040053248A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Hyeeg Inc
/ TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
/ FILE REFERENCE: 784PCT
/ CURRENT APPLICATION NUMBER: US/10/296,115
/ CURRENT FILING DATE: 2002-11-18
/ PRIOR APPLICATION NUMBER: US09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ PRIOR APPLICATION NUMBER: US09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 1478
/ SEQ ID NO 481
/ LENGTH: 1570
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-296-115-481
```

```
Query Match          58.4%; Score 531; DB 17; Length 1570;
Best Local Similarity 99.8%; Pred. No. 3.6e-250;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 328 GTGCTCCAGAAAGCTATAGCTATGACACCTCCAGCTCTTCAAGAGACAGAGGGTAGC 387
DB 76 GTGCTCCAGAAAGCTATAGCTATGACACCTCCAGCTCTTCAAGAGAGAGAGGGTAGC 135
QY 388 TGCCTGCGCGCTGCGAGCTCAAGCAAGTCTGCAAAATTTTCAGAGAGGTGAGAGACA 447
DB 136 TGCCTGCGCGCTGCGAGCTCAAGCAAGTCTGCAAAATTTTCAGAGAGGTGAGAGACA 195
QY 448 GAGTCCCCCAACAAGCGGCGAGCGAGAGTGGAGGCGGCGCCAGTGTGTGCGACA 507
DB 196 GAGTCCCCCAACAAGCGGCGAGCGAGAGTGGAGGCGGCGCCAGTGTGTGCGACA 255
QY 508 CGGGCGGAGAGAGGGGCGCGAGCGGAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 567
DB 256 CGGGCGGAGAGAGGGGCGCGAGCGGAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 315
QY 568 TCTGAAGGCAAAAGTGAAGTGTGACATCCGAGCTCCGGGTTTCAGAGAGTATGCGAGAT 627
DB 316 TCTGAAGGCAAAAGTGAAGTGTGACATCCGAGCTCCGGGTTTCAGAGAGTATGCGAGAT 375
QY 628 GAGCGAGCTTTGAGAGCGGCGGTGCAATCCGGCGGCGCCAGGCGCTGCGCGGAGCTG 687
DB 376 GAGCGAGCTTTGAGAGCGGCGGTGCAATCCGGCGGCGCCAGGCGCTGCGCGGAGCTG 435
QY 688 GAGCTGTTGGAGAGCGGCGAGTGTGCGCTCAAGGAGCTTGGGCTCTGTGTTGT 747
DB 436 GAGCTGTTGGAGAGCGGCGAGTGTGCGCTCAAGGAGCTTGGGCTCTGTGTTGT 495
QY 748 GACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGAGGCGACTAAGTGTGCG 807
DB 496 GACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGAGGCGACTAAGTGTGCG 555
QY 808 GCCCTGCTGAGAGCGGCGGCGGTGTTCTGATCTGAGAGGCGCTTGGAGAGGCTGTGGC 867
DB 556 GCCCTGCTGAGAGCGGCGGCGGTGTTCTGATCTGAGAGGCGCTTGGAGAGGCTGTGGC 615
QY 868 CGGAGAGCTGTTGCGCTGCTGAGTCAAGTGTGAGAGGCTGAC 909
DB 616 CGGAGAGCTGTTGCGCTGCTGAGTCAAGTGTGAGAGGCTGAC 657
```

```
RESULT 9
US-09-799-777-103
/ Sequence 103, Application US/09799777
/ Patent No. US20020091244A1
/ GENERAL INFORMATION:
/ APPLICANT: Lal, Preeti
/ Hillman, Jennifer L.
/ Corley, Neil C.
/ Guejler, Karl J.
/ Baugh, Mariah
/ Sather, Susan
```


Db 571 AATTCTCAGCAGGG 584

RESULT 11

```
US-10-117-722-853
; Sequence 853, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_genes Version 1.0
; SEQ ID NO 853
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(984)
US-10-117-722-853
```

Query Match 30.9%; Score 281; DB 17; Length 1067;
Best Local Similarity 99.3%; Pred. No. 1.5e-127;
Matches 431; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ATGAGCGTATCCGGGATCGAAGCCCGGCGCGTGTGGGAGAGAGATGAGTCTCGAGTAC 60
DB 151 ATGAGCGTATCCGGGATCGAAGCCCGGCGCGTGTGGGAGAGAGATGAGTCTCGAGTAC 210
QY 61 TACGGGATGCTGTGCTTCAACGATATGTTGAGGTGTGGCGGCAACTGACCGAGTGC 120
DB 211 TACGGGATGCTGTGCTTCAACGATATGTTGAGGTGTGGCGGCAACTGACCGAGTGC 270
QY 121 GAGCTGAGAGTCTCTGCGCTTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAAGCC 180
DB 271 GAGCTGAGAGTCTCTGCGCTTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAATCC 330
QY 181 CGGGCCCGCAGCGGAGCTCTGCTGAGCTGAGCGCGCGGCGGAGTGTGCGGAG 240
DB 331 CGGGCCCGCAGCGGAGCTCTGCTGAGCTGAGCGCGCGGCGGAGTGTGCGGAG 390
QY 241 AGCAACTGTGCGTGTGGGCAACTCTGCGGTGTGGCGCGGCAAGCACTGTGCGG 300
DB 391 AGCAACTGTGCGTGTGGGCAACTCTGCGGTGTGGCGCGGCAAGCACTGTGCGG 450
QY 301 CACCTGCGCGGAGCGGCGCGCGGCGGAGTGTCTTCAAGAAAGCTATGCTATGCACTCC 360
DB 451 CACCTGCGCGGAGCGGCGCGCGGCGGAGTGTCTTCAAGAAAGCTATGCTATGCACTCC 510
QY 361 AGCTCTTCAAGAGAGAGAGGAGTGTGCGGTGCGGAGTCAAGCAAGCTTCTGCA 420
DB 511 AGCTCTTCAAGAGAGAGAGGAGTGTGCGGTGCGGAGTCAAGCAAGCTTCTGCA 570
QY 421 AATTCTCAGCAGGG 434
DB 571 AATTCTCAGCAGGG 584
```

RESULT 12

US-10-723-860-1627/c
; Sequence 1627, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:

```
; APPLICANT: Aziz, Natascha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1627
; LENGTH: 111084
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1627
```

Query Match 30.5%; Score 277; DB 18; Length 111084;
Best Local Similarity 99.7%; Pred. No. 6.1e-126;
Matches 327; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAGCGTATCCGGGATCGAAGCCCGGCGCGTGTGGGAGAGAGATGAGTCTCGAGTAC 60
DB 8862 ATGAGCGTATCCGGGATCGAAGCCCGGCGCGTGTGGGAGAGAGATGAGTCTCGAGTAC 8803
QY 61 TACGGGATGCTGTGCTTCAACGATATGTTGAGGTGTGGCGGCAACTGACCGAGTGC 120
DB 8802 TACGGGATGCTGTGCTTCAACGATATGTTGAGGTGTGGCGGCAACTGACCGAGTGC 8743
QY 121 GAGCTGAGAGTCTCTGCGCTTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAAGCC 180
DB 8742 GAGCTGAGAGTCTCTGCGCTTTCTGCTGATGAGGCTCTTGGCGCGCGGAGGCTTAAGCC 8683
QY 181 CGGGCCCGCAGCGGAGCTCTGCTGAGCTGAGCGCGCGGCGGAGTGTGCGGAG 240
DB 8682 CGGGCCCGCAGCGGAGCTCTGCTGAGCTGAGCGCGCGGCGGAGTGTGCGGAG 8623
QY 241 AGCAACTGTGCGTGTGGGCAACTCTGCGGTGTGGCGCGGCAAGCACTGTGCGG 300
DB 8622 AGCAACTGTGCGTGTGGGCAACTCTGCGGTGTGGCGCGGCAAGCACTGTGCGG 8563
QY 301 CACCTGCGCGGAGCGGCGCGCGGCGGAGTGTCTTCAAGAAAGCTATGCTATGCACTCC 328
DB 8562 CACCTGCGCGGAGCGGCGCGCGGCGGAGTGTCTTCAAGAAAGCTATGCTATGCACTCC 8535
```

RESULT 13

US-10-001-254-7

```
; Sequence 7, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Stenner-Liewen, Frank
; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
; FILE REFERENCE: P-12 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
/ SEQ ID NO 18768
/ LENGTH: 838
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 100, 101, 125, 126, 128, 129, 131, 138, 164, 174, 175, 176,
/ LOCATION: 184, 185
/ OTHER INFORMATION: n = A,T,C or G
US-09-814-353-18768
```

```
Query Match      2.4%; Score 22; DB 10; Length 838;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      35 GCGAGCGAGTGCCTCGA 56
DB      381 GCGAGGAGATGAGTCTGGA 360
```

```
RESULT 17
US-09-814-353-13540
/ Sequence 13540, Application US/09814353
/ Publication No. US20030165831A1
/ GENERAL INFORMATION:
/ APPLICANT: Lee, John
/ APPLICANT: Thompson, Pamela
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ FILE REFERENCE: MRI-006B
/ CURRENT APPLICATION NUMBER: US/09/814,353
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US 60/191,031
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: US 60/207,124
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 60/211,940
/ PRIOR FILING DATE: 2000-06-15
/ PRIOR APPLICATION NUMBER: US 60/216,820
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: US 60/220,661
/ PRIOR FILING DATE: 2000-07-25
/ PRIOR APPLICATION NUMBER: US 60/257,672
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 22037
/ SOFTWARE: FastSeq For Windows Version 4.0
/ SEQ ID NO 13540
/ LENGTH: 323
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-814-353-13540
```

```
Query Match      2.3%; Score 21; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      582 GACCTGGACATCGGCTCCG 602
DB      279 GACCTGGACATCGGCTCCG 299
```

```
RESULT 18
US-10-425-114-827/C
/ Sequence 827, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
```

```
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 827
/ LENGTH: 1439
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700150477_FLI
US-10-425-114-827
```

```
Query Match      2.3%; Score 21; DB 17; Length 1439;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      471 GCGCGGAGTCGGCGCGCC 491
DB      171 GCGCGGAGTCGGCGCGCC 151
```

```
RESULT 19
US-10-425-115-1414/C
/ Sequence 1414, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(5322)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 1414
/ LENGTH: 1496
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_11288C.1
US-10-425-115-1414
```

```
Query Match      2.3%; Score 21; DB 18; Length 1496;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      471 GCGCGGAGTCGGCGCGCC 491
DB      173 GCGCGGAGTCGGCGCGCC 153
```

```
RESULT 20
US-10-425-115-46300/C
/ Sequence 46300, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(5322)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 46300
/ LENGTH: 411
```



```

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: clone ID: MRT4577_142231C.1
US-10-425-115-46300

Query Match      2.2%; Score 20; DB 18; Length 411;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      288 CGACCTGCTGCGCGACCTGG 307
DB      154 CGACCTGCTGCGCGACCTGG 135

RESULT 21
US-10-437-963-90512/c
; Sequence 90512, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 90512
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: clone ID: PAT_MRT4530_89175C.1
US-10-437-963-90512

Query Match      2.2%; Score 20; DB 18; Length 424;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      506 GACGCGCGCGGAGAGGGGCC 525
DB      101 GACGCGCGCGGAGAGGGGCC 82

RESULT 22
US-10-156-761-6280/c
; Sequence 6280, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6280
```

```

; LENGTH: 2304
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2304)
US-10-156-761-6280

Query Match      2.2%; Score 20; DB 15; Length 2304;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      221 GCCGCGGCGCAGTGGCGGCGAG 240
DB      2258 GCCGCGGCGCAGTGGCGGCGAG 2239

RESULT 23
US-10-246-330-1
; Sequence 1, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Fah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PASCSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2607)
US-10-246-330-1

Query Match      2.2%; Score 20; DB 16; Length 2610;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      308 CGCGCAAGCGCGCGCGGCGCA 327
DB      1208 CGCGCAAGCGCGCGCGGCGCA 1227

RESULT 24
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
```

```

; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 GCCGCGGCGCAGTGGCGCGAG 240
Db 7567115 GCCGCGGCGCAGTGGCGCGAG 7567134

RESULT 25
US-10-719-993-13899/c
; Sequence 13899, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CU001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13899
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-13899

Query Match
Best Local Similarity 100.0%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTTTCAAGAGACACGA 380
Db 33 GCTTTCAAGAGACACAGA 15

RESULT 26
US-10-027-632-185402
; Sequence 185402, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
```

```

; SEQ ID NO 185402
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-185402

Query Match
Best Local Similarity 100.0%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GTGCTGGAGAGAGATGAG 48
Db 332 GTGCTGGAGAGAGATGAG 350

RESULT 27
US-10-027-632-185403
; Sequence 185403, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 185403
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-185403

Query Match
Best Local Similarity 100.0%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GTGCTGGAGAGAGATGAG 48
Db 332 GTGCTGGAGAGAGATGAG 350

RESULT 28
US-10-027-632-185402
; Sequence 185402, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; SOFTWARE: FASTSEQ for Windows Version 4.0
```

```

; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185402
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-185402

Query Match
Best Local Similarity 100.0%; Score 19; DB 17; Length 625;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GTGCTGGAGAGAGATGAG 48
Db 332 GTGCTGGAGAGAGATGAG 350

RESULT 29
US-10-027-632-185403
; Sequence 185403, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185403
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-185403

Query Match
Best Local Similarity 100.0%; Score 19; DB 17; Length 625;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GTGCTGGAGAGAGATGAG 48
Db 332 GTGCTGGAGAGAGATGAG 350

RESULT 30
US-10-027-632-209406/c
; Sequence 209406, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209407
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-209406

Query Match
Best Local Similarity 100.0%; Score 19; DB 13; Length 682;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 746 GTGACATCAAGTTCTCAGA 764
Db 605 GTGACATCAAGTTCTCAGA 587

RESULT 31
US-10-027-632-209407/c
; Sequence 209407, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209407
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-209407

Query Match
Best Local Similarity 100.0%; Score 19; DB 13; Length 682;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 746 GTGACATCAAGTTCTCAGA 764
| | | | | | | | | | | | | | | | | | | | | |
Db 605 GTGACATCAAGTTCTCAGA 587

RESULT 32
US-10-027-632-209406/c
; Sequence 209406, Application US/10027632
; Publication No. US2003020407SA9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209406
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-209406

Query Match
Best Local Similarity 100.0%; Score 19; DB 17; Length 682;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 746 GTGACATCAAGTTCTCAGA 764
| | | | | | | | | | | | | | | | | | | | | |
Db 605 GTGACATCAAGTTCTCAGA 587

RESULT 33
US-10-027-632-209407/c
; Sequence 209407, Application US/10027632
; Publication No. US2003020407SA9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209407
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-209407

Query Match
Best Local Similarity 100.0%; Score 19; DB 17; Length 682;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 746 GTGACATCAAGTTCTCAGA 764
| | | | | | | | | | | | | | | | | | | | | |
Db 605 GTGACATCAAGTTCTCAGA 587

RESULT 34
US-10-245-752-111
; Sequence 111, Application US/10245752
; Publication No. US20030064473A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Ausetia
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Pong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C66
; CURRENT APPLICATION NUMBER: US/10/245,752
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-752-111

Query Match
Best Local Similarity 100.0%; Score 19; DB 14; Length 943;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GCCCCGTGCTGGAGAGAG 43
| | | | | | | | | | | | | | | | | | | | | |


```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-107-111

Query Match
Best Local Similarity 100.0%; Score 19; DB 14; Length 943;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GCGCCGCTGCTGGAGGAGG 43
DB 437 GCGCCGCTGCTGGAGGAGG 455

RESULT 38
US-10-245-143-111
; Sequence 111, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guirney, Austria
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C90
; CURRENT APPLICATION NUMBER: US/10/245,143
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-107-111
```

```

; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-143-111

Query Match
Best Local Similarity 100.0%; Score 19; DB 14; Length 943;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GCGCCGCTGCTGGAGGAGG 43
DB 437 GCGCCGCTGCTGGAGGAGG 455

RESULT 39
US-10-245-771-111
; Sequence 111, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guirney, Austria
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C98
; CURRENT APPLICATION NUMBER: US/10/245,771
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-771-111

Query Match
Best Local Similarity 100.0%; Score 19; DB 14; Length 943;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 25 GCGCCGCTGCTGGAGGAGG 43
Db 437 GCGCCGCTGCTGGAGGAGG 455

RESULT 40
US-10-245-851-111
; Sequence 111, Application US/10245851
; Publication No. US20030068782A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guiney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Matanbe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C93
; CURRENT APPLICATION NUMBER: US/10/245,851
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-851-111

Query Match 2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GCGCCGCTGCTGGAGGAGG 43
Db 437 GCGCCGCTGCTGGAGGAGG 455

RESULT 41
US-10-245-883-111
; Sequence 111, Application US/10245883
; Publication No. US20030068783A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen

APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Matanbe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C70
; CURRENT APPLICATION NUMBER: US/10/245,883
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-883-111

Query Match 2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GCGCCGCTGCTGGAGGAGG 43
Db 437 GCGCCGCTGCTGGAGGAGG 455

RESULT 42
US-10-237-535-111
; Sequence 111, Application US/10237535
; Publication No. US20030073188A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guiney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Matanbe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C3
; CURRENT APPLICATION NUMBER: US/10/237,535
; CURRENT FILING DATE: 2002-09-06

PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127312
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327

PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206310
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229886
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25


```

; PRIOR APPLICATION NUMBER: 09/380138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/403297
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 09/423741
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: 09/709238
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/802706
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 09/872035
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09
```

```

Query Match          2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      25 GCCCGCTGCTGGAGGAGG 43
          |||
Db       437 GCCCGCTGCTGGAGGAGG 455
```

```

RESULT 43
US-10-238-183-111
; Sequence 111, Application US/10238183
; Publication No. US20030073189A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Zhang, William
; APPLICANT: Zhong, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630RIC11
; CURRENT APPLICATION NUMBER: US/10/238,183
; PRIOR APPLICATION NUMBER: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
```

```

; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091358
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/106932
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/115554
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119342
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/123957
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/127372
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/131271
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/135725
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135729
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138385
; PRIOR FILING DATE: 1999-06-09
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144732
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144790
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145228
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146843
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/148188
; PRIOR FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/148513
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/149327
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149395
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/150114
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/151700
; PRIOR FILING DATE: 1999-08-31
```

PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238

PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 09/872035
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/924419
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 09/927796
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/929404
PRIOR FILING DATE: 2001-08-13
PRIOR APPLICATION NUMBER: 09/931836
PRIOR FILING DATE: 2001-08-16
PRIOR APPLICATION NUMBER: 09/941992
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: 09/946374
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 10/001054
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 10/081056
PRIOR FILING DATE: 2002-02-20
PRIOR APPLICATION NUMBER: 10/119480
PRIOR FILING DATE: 2002-04-09

Query Match 2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GCCCGCTGCTGGAGGAGG 43
Db 437 GCCCGCTGCTGGAGGAGG 455

RESULT 44

US-10-238-283-111
Sequence 111, Application US/10238283
Publication No. US20030073190A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin
APPLICANT: Baton, Dan
APPLICANT: Filvarcoff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C15
CURRENT APPLICATION NUMBER: US/10/238,283
CURRENT FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02

```

; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-238-283-111
```

```

Query Match          2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      25 GCGCCGCTGCTGGAGGAGG 43
         |||||
Db       437 GCGCCGCTGCTGGAGGAGG 455
```

RESULT 45

US-10-238-370-111

```

; Sequence 111, Application US/10238370
; Publication No. US20030073191A1
```

GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guiney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Matambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C10
; CURRENT APPLICATION NUMBER: US/10/238,370
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-238-370-111
```

```

Query Match          2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      25 GCGCCGCTGCTGGAGGAGG 43
         |||||
Db       437 GCGCCGCTGCTGGAGGAGG 455
```

RESULT 46

US-10-245-055-111

```

; Sequence 111, Application US/10245055
; Publication No. US20030073192A1
```

GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin
; APPLICANT: Baton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Guiney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Matambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C8
; CURRENT APPLICATION NUMBER: US/10/245,055
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-245-055-111
```

```

Query Match          2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      25 GCGCCGCTGCTGGAGGAGG 43
         |||||
Db       437 GCGCCGCTGCTGGAGGAGG 455
```

RESULT 47

US-10-245-147-111

```

; Sequence 111, Application US/10245147
; Publication No. US20030073193A1
```

GENERAL INFORMATION:

```
/ APPLICANT: Baker, Kevin
/ APPLICANT: Baton, Dan
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stephan, Jean-Philippe
/ APPLICANT: Watande, Colin
/ APPLICANT: Wood, William
/ APPLICANT: Zhang, Zemin
/ APPLICANT: Fong, Sherman
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3630R1C72
/ CURRENT APPLICATION NUMBER: US/10/245,147
/ PRIOR FILING DATE: 2002-09-16
/ PRIOR APPLICATION NUMBER: 10/197942
/ PRIOR FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/059114
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/063046
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/065027
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/079689
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/086478
/ PRIOR FILING DATE: 1998-05-22
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090689
/ PRIOR FILING DATE: 1998-06-25
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 116
/ SEQ ID NO 111
/ LENGTH: 943
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ US-10-245-147-111

Query Match      2.1% Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
/ FILE REFERENCE: P3630R1C85
/ CURRENT APPLICATION NUMBER: US/10/245,730
/ PRIOR FILING DATE: 2002-09-16
/ PRIOR APPLICATION NUMBER: 10/197942
/ PRIOR FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/059114
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/063046
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/065027
/ PRIOR FILING DATE: 1997-11-10
/ PRIOR APPLICATION NUMBER: 60/079689
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/086478
/ PRIOR FILING DATE: 1998-05-22
/ PRIOR APPLICATION NUMBER: 60/087607
/ PRIOR FILING DATE: 1998-06-02
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090689
/ PRIOR FILING DATE: 1998-06-25
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 116
/ SEQ ID NO 111
/ LENGTH: 943
/ TYPE: DNA
/ ORGANISM: Homo Sapien
/ US-10-245-730-111

Query Match      2.1% Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-739-111

Query Match      2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 GCGCCGTGCTGGAGGAGG 43
        |||||
Db      437 GCGCCGTGCTGGAGGAGG 455

RESULT 50
US-10-246-210-111
; Sequence 111, Application US/10246210
; Publication No. US20030073196A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Bacon, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watande, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C121
; CURRENT APPLICATION NUMBER: US/10/246,210
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 111
; LENGTH: 943
; TYPE: DNA
```

```
; ORGANISM: Homo Sapien
US-10-246-210-111

Query Match      2.1%; Score 19; DB 14; Length 943;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      25 GCGCCGTGCTGGAGGAGG 43
        |||||
Db      437 GCGCCGTGCTGGAGGAGG 455

Search completed: March 24, 2005, 07:34:22
Job time : 833.188 secs
```

THIS PAGE LEFT BLANK

RESULT 4
US-09-252-991A-2875
; Sequence 2875, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2875
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2875

Query Match 2.1%; Score 20; DB 4; Length 2814;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 CGCGCAAGCGCGCGCGCA 327
Db 1371 CGCGCAAGCGCGCGCGCA 1390

RESULT 5
US-09-949-016-21931/c
; Sequence 21931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21931
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21931

Query Match 2.1%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTCAAGAGACAGA 380
Db 584 GCTCTCAAGAGACAGA 566

RESULT 6
US-09-949-016-21932/c
; Sequence 21932, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21932
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21932

Query Match 2.1%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTCAAGAGACAGA 380
Db 233 GCTCTCAAGAGACAGA 215

RESULT 7
US-09-949-016-21933/c
; Sequence 21933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21933
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21933

Query Match 2.1%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTCAAGAGACAGA 380
Db 93 GCTCTCAAGAGACAGA 75

RESULT 8
US-09-949-016-21934/c
; Sequence 21934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21934

Query Match
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAAGAGACAGA 380
Db 77 GCTCTTCAAGAGACAGA 59

RESULT 9
US-09-949-016-21935/c
; Sequence 21935, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21935
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-21935

Query Match
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAAGAGACAGA 380
Db 30 GCTCTTCAAGAGACAGA 12

RESULT 10
US-09-949-016-84252/c
; Sequence 84252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84252
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-84252

Query Match
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAAGAGACAGA 380
Db 584 GCTCTTCAAGAGACAGA 566

RESULT 11
US-09-949-016-84253/c
; Sequence 84253, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84253
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-84253

Query Match
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAAGAGACAGA 380
Db 233 GCTCTTCAAGAGACAGA 215

RESULT 12
US-09-949-016-84254/c
; Sequence 84254, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84254
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
```

US-09-949-016-84254

Query Match 2.1%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAGAGAGACAGA 380
DB 93 GCTCTTCAGAGAGACAGA 75

RESULT 13

US-09-949-016-84255/c
; Sequence 84255, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CU001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 84255
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-84255

Query Match 2.1%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAGAGAGACAGA 380
DB 77 GCTCTTCAGAGAGACAGA 59

RESULT 14

US-09-949-016-84256/c
; Sequence 84256, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CU001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 84256
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-84256

Query Match 2.1%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTTCAGAGAGACAGA 380
DB 30 GCTCTTCAGAGAGACAGA 12

RESULT 15

US-09-270-767-27566
; Sequence 27566, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27566
; LENGTH: 878
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-27566

Query Match 2.1%; Score 19; DB 4; Length 878;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 GCTGAGCTCTGACCTTT 141
DB 388 GCTGAGCTCTGACCTTT 406

RESULT 16

US-09-270-767-11903
; Sequence 11903, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11903
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11903

Query Match 2.1%; Score 19; DB 4; Length 988;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 GCTGAGCTCTGACCTTT 141
DB 498 GCTGAGCTCTGACCTTT 516

RESULT 17

US-09-799-451-383
; Sequence 383, Application US/0799451
; Patent No. 6783869
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.

```

; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Dirmac, Radoje T.
; TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: PL_FL_genes Version 2.0
; SEQ ID NO 383
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1178)
US-09-799-451-383
```

```

Query Match      2.1%; Score 19; DB 4; Length 1742;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      25 GCGCGGCTGCTGGAGGAGG 43
Db      891 GCGCGGCTGCTGGAGGAGG 909
```

```

RESULT 18
US-09-252-991A-5836/C
; Sequence 5836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5836
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5836
```

```

Query Match      2.1%; Score 19; DB 4; Length 2019;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      204 CTGCTGGAGCTGGAGCGC 222
Db      1816 CTGCTGGAGCTGGAGCGC 1798
```

```

RESULT 19
US-09-394-142B-5
; Sequence 5, Application US/09394142B
; Patent No. 6696558
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Takayama, Shinichi
```

```

; TITLE OF INVENTION: No. 6696558e1 BAG Proteins and Nucleic Acid Molecules Encoding
; FILE REFERENCE: P-LJ 3737
; CURRENT APPLICATION NUMBER: US/09/394,142B
; CURRENT FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 09/150,489
; PRIOR FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 2528
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2031)
; NAME/KEY: misc feature
; LOCATION: (1)..(2528)
; OTHER INFORMATION: n= a, c, t or g
US-09-394-142B-5
```

```

Query Match      2.1%; Score 19; DB 4; Length 2528;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      508 CGGCGCGGAGAGGGGCC 526
Db      207 CGGCGCGGAGAGGGGCC 225
```

```

RESULT 20
US-09-394-142B-19
; Sequence 19, Application US/09394142B
; Patent No. 6696558
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Takayama, Shinichi
; TITLE OF INVENTION: No. 6696558e1 BAG Proteins and Nucleic Acid Molecules Encoding
; FILE REFERENCE: P-LJ 3737
; CURRENT APPLICATION NUMBER: US/09/394,142B
; CURRENT FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: 09/150,489
; PRIOR FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 19
; LENGTH: 2534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (307)..(2034)
US-09-394-142B-19
```

```

Query Match      2.1%; Score 19; DB 4; Length 2534;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      508 CGGCGCGGAGAGGGGCC 526
Db      207 CGGCGCGGAGAGGGGCC 225
```

```

RESULT 21
US-09-252-991A-5773
; Sequence 5773, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
```

;; CURRENT APPLICATION NUMBER: US/09/252,991A
;; CURRENT FILING DATE: 1999-02-18
;; PRIOR APPLICATION NUMBER: US 60/074,768
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 5773
;; LENGTH: 2748
;; TYPE: DNA
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5773

Query Match 2.1%; Score 19; DB 4; Length 2748;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 CTTGCTGGAGCTGGAGCGC 222
Db 1086 CTTGCTGGAGCTGGAGCGC 1104

RESULT 22
US-09-902-540-1096/c
;; Sequence 1096, Application US/09902540
;; Patent No. 6833447
;; GENERAL INFORMATION:
;; APPLICANT: Goldman, Barry S.
;; APPLICANT: Hinkle, Gregory J.
;; APPLICANT: Slater, Steven C.
;; APPLICANT: Wiegand, Roger C.
;; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;; FILE REFERENCE: 38-10(15849)B
;; CURRENT APPLICATION NUMBER: US/09/902,540
;; CURRENT FILING DATE: 2001-07-10
;; PRIOR APPLICATION NUMBER: 60/217,883
;; PRIOR FILING DATE: 2000-07-10
;; NUMBER OF SEQ ID NOS: 16825
;; SEQ ID NO 1096
;; LENGTH: 14555
;; TYPE: DNA
;; ORGANISM: Myxococcus xanthus
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1) ..(14555)
;; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1096

Query Match 2.1%; Score 19; DB 4; Length 14555;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 204 CTTGCTGGAGCTGGAGCGC 222
Db 2734 CTTGCTGGAGCTGGAGCGC 2716

RESULT 23
US-09-949-016-16766/c
;; Sequence 16766, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; FILE REFERENCE: CLO01307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 16766
;; LENGTH: 32584
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-16766

Query Match 2.1%; Score 19; DB 4; Length 32584;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 729 CTTGCTGGAGCTGGAGCGC 747
Db 9895 CTTGCTGGAGCTGGAGCGC 9877

RESULT 24
US-09-949-016-12419
;; Sequence 12419, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; FILE REFERENCE: CLO01307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12419
;; LENGTH: 60095
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-12419

Query Match 2.1%; Score 19; DB 4; Length 60095;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 CTTTCTGTGATGAGGCT 156
Db 9576 CTTTCTGTGATGAGGCT 9594

RESULT 25
US-09-949-016-14457
;; Sequence 14457, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; FILE REFERENCE: CLO01307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14457
;; LENGTH: 60095

```
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14457

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 60095;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 CTTTCTGCTGATGAGGCT 156
Db 9576 CTTTCTGCTGATGAGGCT 9594

RESULT 26
US-09-949-016-14157/c
; Sequence 14157, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14157
; LENGTH: 455726
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(455726)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-14157

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 455726;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTCAAGAGGACAGA 380
Db 410955 GCTCTCAAGAGGACAGA 410937

RESULT 27
US-09-949-016-11940/c
; Sequence 11940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11940
; LENGTH: 481115
; TYPE: DNA
```

```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(481115)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-11940

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 481115;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 362 GCTCTCAAGAGGACAGA 380
Db 388344 GCTCTCAAGAGGACAGA 388326

RESULT 28
US-09-513-999C-34541
; Sequence 34541, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 34541
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 79
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 80
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 288
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 289
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 361
; OTHER INFORMATION: s=g or c
; US-09-513-999C-34541

Query Match
Best Local Similarity 100.0%; Score 18; DB 4; Length 397;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 516 GAGAGGGGCCCCAGCGC 533
Db 82 GAGAGGGGCCCCAGCGC 99

RESULT 29
US-09-502-540-6231
; Sequence 6231, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
```

```

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6231
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-6231

Query Match      2.0%; Score 18; DB 4; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      210 GGAGCTGAGCGCGCGG 227
Db      195 GGAGCTGAGCGCGCGG 212

RESULT 30
US-09-949-016-33681/C
; Sequence 33681, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33681
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-33681

Query Match      2.0%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      714 GCTGCGCTCAAGGACCT 731
Db      463 GCTGCGCTCAAGGACCT 446

RESULT 31
US-09-949-016-69186/C
; Sequence 69186, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69186
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-69186

Query Match      2.0%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      714 GCTGCGCTCAAGGACCT 731
Db      463 GCTGCGCTCAAGGACCT 446

RESULT 32
US-09-902-540-8543
; Sequence 8543, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8543
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8543

Query Match      2.0%; Score 18; DB 4; Length 969;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      302 ACCTGGCGCGCAAGCGC 319
Db      719 ACCTGGCGCGCAAGCGC 736

RESULT 33
US-09-252-991A-13074
; Sequence 13074, Application US/0925291A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13074
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
FEATURE:
```

NAME/KEY: unsure
LOCATION: (138)
OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-13074

Query Match 2.0%; Score 18; DB 4; Length 1011;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 CGACCTGCTGCGCAGCT 305
DB 510 CGACCTGCTGCGCAGCT 527

RESULT 34
US-09-363-189B-3
Sequence 3, Application US/09363189B
Patent No. 6242228
GENERAL INFORMATION:
APPLICANT: SUGIYAMA, MASAKAZU
APPLICANT: TONOUCHI, MOTO
APPLICANT: SUZUKI, SHUNICHI
APPLICANT: YOKOZAKI, KENZO
TITLE OF INVENTION: XYLTOL DEHYDROGENASE OF ACETIC ACID BACTERIA AND GENE THEREOF
FILE REFERENCE: 0010-1024-0
CURRENT APPLICATION NUMBER: US/09/363,189B
CURRENT FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: JP10-216047
PRIOR FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.0
SEQ ID NO 3
LENGTH: 1056
TYPE: DNA
ORGANISM: Gluconobacter oxydans
FEATURE:
NAME/KEY: CDS
LOCATION: (25)..(1053)
US-09-363-189B-3

Query Match 2.0%; Score 18; DB 3; Length 1056;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 CGTGGCGTGGCAGTCA 408
DB 237 CGTGGCGTGGCAGTCA 254

RESULT 35
US-09-902-540-9670/C
Sequence 9670, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 9670
LENGTH: 1242
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1242)
OTHER INFORMATION: unsure at all n locations

US-09-902-540-9670

Query Match 2.0%; Score 18; DB 4; Length 1242;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 GCTGAGCGCGCGGCA 230
DB 289 GCTGAGCGCGCGGCA 272

RESULT 36
US-09-902-540-8471
Sequence 8471, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 8471
LENGTH: 1268
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-8471

Query Match 2.0%; Score 18; DB 4; Length 1268;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CTGCTGAGCTGAGCGC 222
DB 313 CTGCTGAGCTGAGCGC 330

RESULT 37
US-09-902-540-5865/C
Sequence 5865, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 5865
LENGTH: 1294
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-5865

Query Match 2.0%; Score 18; DB 4; Length 1294;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 526 CAGCGCAGCCAGCAG 543
DB 301 CAGCGCAGCCAGCAG 284

RESULT 38
US-09-902-540-149/C
; Sequence 149, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 149
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-149

Query Match 2.0%; Score 18; DB 4; Length 1296;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 526 CCAGCCGACCCCGACGAG 543
Db 303 CCAGCCGACCCCGACGAG 286

RESULT 39
US-09-252-991A-10421/C
; Sequence 10421, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10421
; LENGTH: 1419
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10421

Query Match 2.0%; Score 18; DB 4; Length 1419;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 10 TCCGGGTGACCCCGGCC 27
Db 1259 TCCGGGTGACCCCGGCC 1242

RESULT 40
US-09-252-991A-10319/C
; Sequence 10319, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10319
; LENGTH: 1695
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10319

Query Match 2.0%; Score 18; DB 4; Length 1695;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 10 TCCGGGTGACCCCGGCC 27
Db 1506 TCCGGGTGACCCCGGCC 1489

RESULT 41
US-09-252-991A-12867/C
; Sequence 12867, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12867
; LENGTH: 1809
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1486)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-252-991A-12867

Query Match 2.0%; Score 18; DB 4; Length 1809;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 288 CGACCTGCTGCCGACCT 305
Db 1114 CGACCTGCTGCCGACCT 1097

RESULT 42
US-09-799-451-868
; Sequence 868, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aiqing J.
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yundong
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei

```
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 868
; LENGTH: 1899
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (314)..(1174)
US-09-799-451-868
```

```
Query Match          2.0%; Score 18; DB 4; Length 1899;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      839 TGA CTGAGGCGCTGCGAG 856
Db       1178 TGA CTGAGGCGCTGCGAG 1195
```

```
RESULT 43
US-09-902-540-335/c
; Sequence 335, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 335
; LENGTH: 2189
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-335
```

```
Query Match          2.0%; Score 18; DB 4; Length 2189;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      210 GGA GCTGAGGCGCGCG 227
Db       1995 GGA GCTGAGGCGCGCGCG 1978
```

```
RESULT 44
US-09-252-991A-10741
; Sequence 10741, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10741
; LENGTH: 2376
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10741
```

```
Query Match          2.0%; Score 18; DB 4; Length 2376;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      10 TCCGGGTGACCCCGGCC 27
Db       200 TCCGGGTGACCCCGGCC 217
```

```
RESULT 45
US-09-799-451-66
; Sequence 66, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungang
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 66
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8)..(2530)
US-09-799-451-66
```

```
Query Match          2.0%; Score 18; DB 4; Length 3016;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      709 GCA GTGCTGCGCTCAAG 726
Db       1827 GCA GTGCTGCGCTCAAG 1844
```

```
RESULT 46
US-09-949-016-12546
; Sequence 12546, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12546
; LENGTH: 5325
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12546

Query Match          2.0%; Score 18; DB 4; Length 5325;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      714 GCTGCGCTCAAGGAGCCT 731
      |||
Db      1839 GCTGCGCTCAAGGAGCCT 1856

RESULT 47
US-09-949-016-13773
; Sequence 13773, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13773
; LENGTH: 5326
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13773

Query Match          2.0%; Score 18; DB 4; Length 5326;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      714 GCTGCGCTCAAGGAGCCT 731
      |||
Db      1839 GCTGCGCTCAAGGAGCCT 1856

RESULT 48
US-09-902-540-891
; Sequence 891, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
```

```

; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 891
; LENGTH: 6380
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-891

Query Match          2.0%; Score 18; DB 4; Length 6380;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      205 CTGCTGAGCTGAGCGGC 222
      |||
Db      5425 CTGCTGAGCTGAGCGGC 5442

RESULT 49
US-09-902-540-900/c
; Sequence 900, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 900
; LENGTH: 7713
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-900

Query Match          2.0%; Score 18; DB 4; Length 7713;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      302 ACTTGCGCGCGCAGCGGC 319
      |||
Db      6994 ACTTGCGCGCGCAGCGGC 6977

RESULT 50
US-09-949-016-17210/c
; Sequence 17210, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17210
; LENGTH: 8011
; TYPE: DNA
```

; ORGANISM: Human
US-09-949-016-17210

Query Match 2.0%; Score 18; DB 4; Length 8011;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 440 GGAGACAGGCTCCCCC 457
|||||
Db 4226 GGAGACAGGCTCCCCC 4209

Search completed: March 24, 2005, 03:14:48
Job time : 169.089 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 01:57:52 ; Search time 67.1055 Seconds
(without alignments)
5399.743 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 620
Sequence: 1 aggcgcataatagagaag.....ctgaaaggcatagtg99 1883

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565918

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+ n2p model -DEV=xlp
-Q/cgr2.1/USPTO.spool/p/US10030271/runat_21032005_153547_26829/app_query.fasta_1.3150
-DB=PIR_79 -Qfmt=fastran -SUBFIX=oligo -rpr -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL -OUTFMT=pco
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10030271 @CGN 1.1.151 @runat 21032005_153547_26829 -NCPUS=6 -ICU=3
-NO_MMAP -LARGEOUTERY -NBS_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6
-Fgapext=7 -Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	1.6	235	2	hypothetical prote
2	10	1.6	1460	1	immediate-early pr
3	10	1.5	127	2	hypothetical prote
4	9	1.5	307	2	hypothetical prote
5	9	1.5	370	1	transforming prote
6	9	1.5	454	2	hypothetical prote
7	9	1.5	614	2	glucose inhibited
8	9	1.3	75	2	ATP synthase C cha
9	9	1.3	75	2	ATP synthase chain
10	8	1.3	134	2	methylinamonyl-CoA
11	8	1.3	158	2	hypothetical prote
12	8	1.3	170	2	hypothetical prote
13	8	1.3	170	2	hypothetical prote
14	8	1.3	195	2	hypothetical prote

15	8	1.3	195	2	S45627	acidic endoprotein
16	8	1.3	197	2	C70601	hypothetical prote
17	8	1.3	206	2	AG0466	probable homoserin
18	8	1.3	213	2	B75525	transcription regu
19	8	1.3	219	2	A57694	transcription fact
20	8	1.3	219	2	G87371	hypothetical prote
21	8	1.3	224	2	B87197	hypothetical prote
22	8	1.3	224	2	E70790	probable transcrip
23	8	1.3	236	1	T46967	diheme cytochrome
24	8	1.3	262	2	C70648	probable nuot prot
25	8	1.3	265	2	AH0775	hydroxyethylthioaz
26	8	1.3	266	2	AE0156	probable transcrip
27	8	1.3	279	2	B70328	hypothetical prote
28	8	1.3	279	2	T34848	probable transcrip
29	8	1.3	282	2	S46793	vacuolar protein s
30	8	1.3	288	2	F48423	homeotic protein e
31	8	1.3	293	2	G87018	probable membrane
32	8	1.3	314	2	C87667	ABC transporter, A
33	8	1.3	339	2	S20880	homeotic protein H
34	8	1.3	342	2	S18649	homeotic protein H
35	8	1.3	347	2	T08826	secretory carrier
36	8	1.3	363	2	T34931	hypothetical prote
37	8	1.3	366	2	T51339	mitogen-activated
38	8	1.3	374	2	AE2227	transposase alr337
39	8	1.3	375	2	T05707	phosphate transpor
40	8	1.3	376	2	C75580	adenine deaminase-
41	8	1.3	381	2	S08296	beta-lactamase (EC
42	8	1.3	381	2	S45109	beta-lactamase (EC
43	8	1.3	381	2	AD0897	conserved hypochet
44	8	1.3	387	2	D88479	protein P47D12.1 (
45	8	1.3	394	2	E87675	hypothetical prote
46	8	1.3	397	2	T08345	hypothetical prote
47	8	1.3	400	1	SYZMW1	naringenin-chalcon
48	8	1.3	402	2	A45056	osteogenic protein
49	8	1.3	406	2	H69812	conserved hypochet
50	8	1.3	410	2	E70475	foliopolyglutamate
51	8	1.3	413	2	AE8726	conserved hypochet
52	8	1.3	417	2	B86858	cell division prot
53	8	1.3	434	2	C83779	aminotransferase B
54	8	1.3	435	2	AG3257	CBS domain contain
55	8	1.3	460	1	SS1516	serine-type carbox
56	8	1.3	460	2	T07868	DNA binding protei
57	8	1.3	463	2	S77558	hypothetical prote
58	8	1.3	474	2	D64855	TRK system potassi
59	8	1.3	475	2	A38340	66K glycoprotein p
60	8	1.3	477	1	ORHUB1	beta-1-adrenergic
61	8	1.3	480	2	I53053	beta-1-adrenergic
62	8	1.3	513	1	S43342	flavonoid 3',5'-hy
63	8	1.3	521	2	F87317	hypothetical prote
64	8	1.3	528	2	AC3236	hypothetical prote
65	8	1.3	568	2	S65770	malicooligosyltreha
66	8	1.3	665	2	G96530	probable acyl CoA
67	8	1.3	666	2	T07928	probable long-chain
68	8	1.3	683	2	C71322	probable translati
69	8	1.3	738	1	TFHUM	melanotransferrin
70	8	1.3	761	2	S20458	pqqf protein - Kle
71	8	1.3	767	2	A35645	major surface prot
72	8	1.3	775	2	I49759	hepatocyte growth
73	8	1.3	786	2	S22155	oncogene 1 (tre-2
74	8	1.3	798	1	B44051	infected cell prot
75	8	1.3	828	1	T33481	hypothetical prote
76	8	1.3	845	2	E72652	pyruvate, water di
77	8	1.3	853	2	T23697	hypothetical prote
78	8	1.3	862	2	T46289	hypothetical prote
79	8	1.3	864	2	T08575	protein kinase hom
80	8	1.3	877	2	T35861	probable large sec
81	8	1.3	914	2	T00757	probable ubiquitin
82	8	1.3	928	2	T04194	hypothetical prote
83	8	1.3	937	2	T09494	hypothetical prote
84	8	1.3	963	2	T09497	ubiquitin cholest
85	8	1.3	1052	1	A44937	kinetoplast-associ
86	8	1.3	1073	2	T38763	hypothetical prote
87	8	1.3	1089	2	S22158	transforming prote

88 8 1.3 1095 2 B83471 probable pyruvate
C 89 8 1.3 1162 2 T51040 hypothetical prote
C 90 8 1.3 1222 2 G59100 hypothetical prote
C 91 8 1.3 1293 2 T30871 orsellinic acid sy
C 92 8 1.3 1704 2 T43141 vitellogenin 1 - m
C 93 8 1.3 1733 1 B45344 probable nuclear a
C 94 8 1.3 1788 2 T31095 vitellogenin precu
C 95 8 1.3 1958 2 B40505 hypothetical prote
C 96 8 1.3 2279 2 T42531 acetyl-CoA carboxy
C 97 8 1.3 2280 2 T38906 acetyl-CoA carboxy
C 98 8 1.3 4848 2 T30289 pristinamycin I sy
C 99 7 1.1 12 2 A60528 insulin-like growt
100 7 1.1 12 2 A33520 inhibitory diffus

ALIGNMENTS

RESULT 1

T30656 hypothetical protein 54L - Molluscum contagiosum virus 1

N:Alternate names: MC054L

C:Species: Molluscum contagiosum virus 1

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C/Accession: T30656

R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.

Science 273, 813-816, 1996

A>Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re

A/Reference number: Z20876; MUID:96325459; PMID:8670425

A/Accession: T30656

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-235 <SEN>

A/Cross-references: UNIPROT:Q98222; EMBL:U60315; PIDN:AA05182.1

C/Genetics:

A/Note: MC054L

Alignment Scores:

Pred. No.: 1.54 Length: 235
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x T30656 (1-235)

Qy 622 GGTGCCAGCGCGCGGAGAGGCGCCCA 651

Db 156 GYLAHARGARGARGARGGLYALPR 165

RESULT 2

immediate-early protein IE180 - suid herpesvirus 1 (strain Indiana-Funkhauser)

C:Species: suid herpesvirus 1

C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004

C/Accession: S04713

R:Cheung, A.K.

Nucleic Acids Res. 17, 4637-4646, 1989

A>Title: DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies vi

A/Reference number: S04713; MUID:89315207; PMID:2546124

A/Accession: S04713

A/Molecule type: DNA

A/Residues: 1-1460 <CHE>

A/Cross-references: UNIPROT:P11675

C/Superfamily: herpesvirus immediate-early protein IE175

C/Keywords: DNA binding; early protein; transcription regulation

Alignment Scores:

Pred. No.: 1.02 Length: 1460
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.62% Indels: 0

DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x EDBE1F (1-1460)

Qy 659 GGTGGCGTGGGAGCCCTCTCCGCGCCGT 630

Db 1047 GYLAHARGARGARGGLYALPR 1056

RESULT 3

A72522 hypothetical protein ABE2150 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C/Accession: A72522

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Maesuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A/Reference number: A72450; MUID:99310339; PMID:10382966

A/Accession: A72522

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-127 <KAW>

A/Cross-references: UNIPROT:Q9Y9Y9; DBJ:AF000063; NID:G510564; PIDN:BA81161.1; PID:dl

A/Experimental source: strain K1

C/Genetics:

A/Note: ABE2150

Alignment Scores:

Pred. No.: 15.9 Length: 127
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x A72522 (1-127)

Qy 469 AGCTATGACCTCTCCAGCTTTCNAAG 495

Db 95 SerTyrGlyThrsSerSerSerLys 103

RESULT 4

C75584 hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: C75584

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; MUID:20036896; PMID:10567266

A/Accession: C75584

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-307 <WHI>

A/Cross-references: UNIPROT:Q9RYL0; GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AA01251

A/Experimental source: strain R1

A/Genetics:

A/Map position: 2

C/Superfamily: Deinococcus radiodurans hypothetical protein DRA0302

Alignment Scores:

Pred. No.: 13 Length: 307
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x C75584 (1-307)

QY 349 CGCGGCGCTCCAGCTCCAGCAGAGCT 323
DB 278 ArgGlyAlaProAlaProAlaGlyAla 286

RESULT 5

transforming protein int-1 - human
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C/Accession: A24674
R:van Ooyen, A.; Kwee, V.; Nusee, R.
EMBO J. 4, 2905-2909, 1985
A>Title: The nucleotide sequence of the human int-1 mammary oncogene; evolutionary conse
A/Reference number: A24674; MUID:86055728; PMID:2998762
A/Accession: A24674
A/Molecule type: DNA
A/Residues: 1-370 <VAN>
C/Genetics:
A/Genes: GDB:MNT1, INT1
A/Cross-references: GDB:120101, OMIM:164820
A/Map position: 12q13-12q13
A/Intons: 35/2; 120/1; 208/3
C/Superfamily: int-1 transforming protein
C/Keywords: proto-oncogene; transforming protein; transmembrane protein
F:148/Domain: transmembrane #status predicted <TM>

Alignment Scores:
Pred. No.: 12.5 Length: 370
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x TVHUT1 (1-370)

QY 366 GCTCTCGCGCACTGCGCGCGCTC 340
DB 18 AlaleuAlaAlaLeuProAlaAlaLeu 26

RESULT 6

hypothetical protein alr1312 [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C/Accession: AE1970
R:Kaneho, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasanoto, S.; Watanabe, A.; Iriuchih
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AE1970
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-454 <KUR>
A/Cross-references: UNIPROT:Q8YXA3; GB:BA000019; PIDN:BA873269.1; PID:g17330659; GSPDB:C
A/Experimental source: strain PCC 7120
C/Genetics:
A/Genes: alr1312
C/Superfamily: conserved hypothetical protein b0835

Alignment Scores:
Pred. No.: 11.9 Length: 454
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x AE1970 (1-454)

QY 651 AGCGGCGCGCGCGCGCGCGCGCG 677
DB 175 SerArgThrProAlaAlaValArgAla 183

RESULT 7

glucose inhibited division protein UN039 [imported] - Ureaplasma urealyticum
C/Species: Ureaplasma urealyticum
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C/Accession: D82942
R:Glass, J.I.; Leftkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A/Reference number: A82870
A/Accession: D82942
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-614 <GLA>
A/Cross-references: GB:AE002103; GB:AF222894; NID:g6898977; PIDN:AAF30444.1; GSPDB:GN001.
A/Experimental source: serovar 3; biovar 1
C/Genetics:
A/Genes: glda; UN039
A/Genetic code: SGC3
C/Superfamily: glda protein

Alignment Scores:
Pred. No.: 11.2 Length: 614
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x D82942 (1-614)

QY 1670 TCCAGAGCTGACATGACTTGACATT 1696
DB 429 SerArgAlaGlnHisArgLeuAlaLeu 437

RESULT 8

ATP synthase C chain atpc [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AE2664
R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McCelli
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.
A>Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE2664
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-75 <KUP>
A/Cross-references: UNIPROT:Q8UH67; GB:AE008688; PIDN:AL41731.1; PID:g17739081; GSPDB:G
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Genes: atpc
A/Map position: circular chromosome
C/Superfamily: H+-transporting ATP synthase lipid-binding protein

Alignment Scores:
Pred. No.: 162 Length: 75
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x AE2664 (1-75)

QY 913 GGGGACTGAGTGGCGCCCTG 936
 |||
 DB 30 GlyAspTyrLeuSerGlyAlaLeu 37

RESULT 9

D97446
 ATP synthase chain C (AF054609) [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C/Species: Agrobacterium tumefaciens
 C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C/Accession: D97446
 R/Gooder, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A/Reference number: A97359; NCID:21608551; PMID:11743194
 A/Accession: D97446
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-75 <KTR>
 A/Cross-references: UNIPROT:Q8UHG7; GB:AE007869; PIDN:AK86525.1; PID:G15155683; GSPDB:G
 A/Gene: AGR_C_1297
 A/Map position: circular chromosome
 C/Superfamily: H+-transporting ATP synthase lipid-binding protein

Alignment Scores:

Pred. No.: 162 Length: 75
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.29% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x D97446 (1-75)

QY 913 GGGGACTGAGTGGCGCCCTG 936
 |||
 DB 30 GlyAspTyrLeuSerGlyAlaLeu 37

RESULT 10

T44984
 methylmalonyl-CoA decarboxylase (EC 4.1.1.41) gamma chain [imported] - Propionigenium mo
 C/Species: Propionigenium modestum
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C/Accession: T44984
 R/Bott, M.; Pfister, K.; Burda, P.; Kalbermatter, O.; Woehle, G.; Dimroth, P.
 Eur. J. Biochem. 250, 590-599, 1997
 A/Title: Methylmalonyl-CoA decarboxylase from Propionigenium modestum: Cloning and seque
 A/Reference number: Z22888; NCID:98088990; PMID:9428714
 A/Accession: T44984
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-134 <BOT>
 A/Cross-references: UNIPROT:Q54030; EMBL:AJ002015; NID:G2706397; PIDN:CAA05139.1; PID:G2
 A/Experimental source: DSM 2376
 C/Comment: Methylmalonyl-CoA decarboxylase catalyses the only energy-conserving step dur
 1-CoA is coupled to the vectorial transport of Na+ across the cytoplasmic membrane, then
 C/Genetics:
 A/Note: mmdC
 C/Complex: heterotetramer [validated, NCID:98088990]
 C/Function:
 A/Description: EC 4.1.1.41 [validated, NCID:98088990]
 A/Note: Specific activity up to 25 U/mg protein; Km value for (S)-methylmalonyl-CoA of 8
 C/Superfamily: biotin carboxyl carrier protein, lipoyl/biotin-binding homology
 C/Keywords: carbon-carbon lyase; carboxy-lyase

Alignment Scores:

Pred. No.: 142 Length: 134
 Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x T44984 (1-134)

QY 814 CGTCCAGCTGCCCGCCAGCGCCT 791
 |||
 DB 32 ArgProAlaAlaAlaProAlaPro 39

RESULT 11

T26444
 hypothetical protein Y113G7B.11 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T26444
 R.Rienhard, N.
 submitted to the EMBL Data Library, September 1999
 A/Reference number: Z20215
 A/Accession: T26444
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-158 <WIL>
 A/Cross-references: UNIPROT:Q9UX23; EMBL:AL110477; NID:el542121; PIDN:CAB54332.1; CESP:Y1
 A/Experimental source: clone Y113G7B
 C/Genetics:
 A/Gene: CESP:Y113G7B.11
 A/Introns: 30/1; 60/1; 99/1

Alignment Scores:

Pred. No.: 137 Length: 158
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x T26444 (1-158)

QY 341 TCCAGCTCCAGCAGAGCTTAGG 318
 |||
 DB 105 SerSerSerSerArgSerSerArg 112

RESULT 12

C84492
 hypothetical protein At2g10550 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C/Accession: C84492
 R/Alin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umeyam, L.; Tallon,
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A/Reference number: A84420; NCID:20083487; PMID:10617197
 A/Accession: C84492
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-170 <STO>
 A/Cross-references: UNIPROT:Q9S187; GB:AE002093; NID:ig4733994; PIDN:AAD28673.1; GSPDB:GN
 C/Genetics:
 A/Gene: At2g10550
 A/Map position: 2

Alignment Scores:

Pred. No.: 135 Length: 170
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.29% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x C84492 (1-170)

QY 841 CGCTCAGGAGCTGCGCTCTGTG 864
Db 37 ArgSerArgSerArgSerArgSerVal 44

RESULT 13

F83325
hypothetical protein PA2562 [imported] - Pseudomonas aeruginosa (strain PA01)
C/Species: Pseudomonas aeruginosa
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: F83325
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Micooguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A/Reference number: A82950; MUID:20437337; PMID:10984043
A/Accession: F83325
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-170 <SNO>
A/Cross-references: UNIPROT:Q91053; GB:AE004684; GB:AE004091; NID:g9948617; PIDN:AA0595
A/Experimental source: strain PA01
C/Genetics:
A/Gene: PA2562

Alignment Scores:

Pred. No.: 135 Length: 170
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x F83325 (1-170)

QY 900 GGAGCGCTCTGCGGCGACTACT 923
Db 130 G1yArgLeuG1yArgLeuPro 137

RESULT 14

E75462
hypothetical protein - Deinococcus radiodurans (strain R1)
C/Species: Deinococcus radiodurans
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: E75462
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A/Reference number: A75250; MUID:20036896; PMID:10567266
A/Accession: E75462
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-195 <WHI>
A/Cross-references: UNIPROT:Q9RVX4; GB:AE001942; GB:AE000513; NID:g6458611; PIDN:AA1047
A/Experimental source: strain R1
C/Genetics:
A/Gene: DR0897
A/Map position: 1
C/Superfamily: Deinococcus radiodurans hypothetical protein DR0897

Alignment Scores:

Pred. No.: 131 Length: 195
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x E75462 (1-195)

QY 222 GGCGGCGCACTGACCGAGTCCA 245
Db 60 G1yArgAlaThrArgValArg 67

RESULT 15

S45627
acidic endoproteinase precursor - Myxococcus xanthus (strain DK101)
C/Species: Myxococcus xanthus
A/Variety: strain DK101
C/Date: 10-Dec-1994 #sequence_revision 26-Apr-1996 #text_change 09-Jul-2004
C/Accession: S45627; S62857
R/Lucas, N.; Mazaud-Aujard, C.; Bremaud, L.; Cenatiempo, Y.; Julien, R.
Eur. J. Biochem. 222, 247-254, 1994
A/Title: Protein purification, gene cloning and sequencing of an acidic endoprotease from
A/Reference number: S45627; MUID:94291618; PMID:8020464
A/Accession: S45627
A/Molecule type: DNA
A/Residues: 1-195 <LUC>
A/Cross-references: UNIPROT:Q50902; EMBL:X75892; NID:g516391; PIDN:CAA53499.1; PID:g51631
A/Experimental source: strain DK101
A/Accession: S62857
A/Molecule type: protein
A/Residues: 65-101 <LUW>
C/Genetics:
A/Gene: Maep
F/1-29/Domain: signal sequence #status predicted <SIG>
F/30-64/Domain: propeptide #status predicted <PRO>
F/65-195/Product: acidic endoproteinase #status experimental <MAT>

Alignment Scores:

Pred. No.: 131 Length: 195
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x S45627 (1-195)

QY 405 CGCGCAGCACTGCTGCGGAGCT 428
Db 150 ProProArgProAlaAlaPro 157

RESULT 16

C70601
hypothetical protein RV092c - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: C70601
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:9825987; PMID:9634230
A/Accession: C70601
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-197 <COL>
A/Cross-references: UNIPROT:Q05575; GB:Z94752; GB:AL123456; NID:g3261731; PIDN:CAB08152.J
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: RV092c
C/Superfamily: human 5-formyltetrahydrofolate cyclo-ligase

Alignment Scores:

Pred. No.: 130 Length: 197
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x C70601 (1-197)

QY 289 GCCGAGGCTTACCGGCGCCGC 312

DB 95 AAGGAGGCTTACCGGCGCCGC 102

RESULT 17

AG0466 probable homoserine/homoserine lactone efflux protein rhlB [imported] - Yersinia pestis

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: AG0466

R/ParKhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.

demo-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davies, P.; Dougan, G.;

Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,

Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AG0466

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-206 <R>

A/Cross-references: UNIPROT:Q8ZAH0; GB:AL590842; PIDN:CAC93299.1; PID:g13981746; GSPDB:G

C/Genetics:

A/Gene: rhlB

Alignment Scores:

Pred. No.: 129 Length: 206

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.30% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x AG0466 (1-206)

QY 387 TTGCCCCAGCAGCGCGGCTTGTCT 364

DB 138 LTPGCGGCGGCGCGGCTTGTCT 145

RESULT 18

E75525 transcription regulator, Tetr family - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: E75525

R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75520; MUID:20036896; PMID:10567266

A/Accession: E75525

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-213 <WHI>

A/Cross-references: UNIPROT:Q9RXD6; GB:AE001898; GB:AE000513; NID:G6458057; PIDN:AAF0995

C/Genetics:

A/Experimental source: strain R1

C/Genetics:

A/Gene: DR0378

A/Map position: 1

Alignment Scores:

Pred. No.: 128 Length: 213

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.30% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x E75525 (1-213)

QY 326 AGCTTAGGCGCGCTGCGGCGCCGC 303

DB 8 SerSerArgProLeuArgGlnArg 15

RESULT 19

A57694 transcription factor IID 30k chain - human

C/Species: Homo sapiens (man)

C/Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004

C/Accession: A57694; A54981

R/Scheer, E.; Mattei, M.G.; Jacq, X.; Chambon, P.; Tora, L.

Genomics 29, 269-272, 1995

A/Title: Organization and chromosomal localization of the gene (TAF2H) encoding the human

A/Reference number: A57694; MUID:96079120; PMID:8530084

A/Accession: A57694

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-219 <SCH>

A/Cross-references: UNIPROT:Q12962; GB:U25816

A/Note: authors fail to translate CAG as Gln for residue 190

R/Jacq, X.; Brou, C.; Lutz, Y.; Davidson, I.; Chambon, P.; Tora, L.

Cell 79, 107-117, 1994

A/Title: Human TAF-II30 is present in a distinct TFIID complex and is required for trans

A/Reference number: A54981; MUID:95007771; PMID:7923369

A/Accession: A54981

A/Status: preliminary; not compared with conceptual translation

A/Residues: 1-189,191-219 <JAC>

A/Molecule type: mRNA

A/Cross-references: GB:U13991; NID:G562076; PIDN:AAA62230.1; PID:G562077

C/Genetics:

A/Gene: GDB:TAF2H; TAF2A

A/Cross-references: GDB:453296; OMIM:600475

A/Map position: 11p15.3-11p15.3

A/Introns: 78/1, 129/3, 151/2, 189/3

Alignment Scores:

Pred. No.: 127 Length: 219

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.29% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x A57694 (1-219)

QY 789 CCAGGCGCTGCGCGCGGCTGGA 812

DB 52 ProGlyAlaGlyAlaAlaAlaGly 59

RESULT 20

G87371 hypothetical protein CC0987 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: G87371

R/Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolton

n, J.; Arnold, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: G87371

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-219 <STO>

A/Cross-references: UNIPROT:Q9A9J3; GB:AE005673; NID:G13422271; PIDN:AAK22971.1; GSPDB:G

C/Genetics:

A/Gene: CC0987

Alignment Scores:

Pred. No.: 127 Length: 219

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x G87371 (1-219)

QY 433 GCGCAGTGGCGCAGCGTGGT 410
 Db 125 AlaProGlyAlaAlaAlaGlyArg 132

RESULT 21

B87197

hypothetical protein ML2302 [imported] - Mycobacterium leprae

C/Species: Mycobacterium leprae
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: B87197
 R/Col: S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 409, 1007-1011, 2001
 A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sg

A/Title: Massive gene decay in the leprosy bacillus.
 A/Reference number: A86909; MUID:21128732; PMID:11234002

A/Accession: B87197
 A/Status: preliminary
 A/Molecule type: DNA

A/Residues: 1-224 <MOD>
 A/Cross-references: UNIPROT:Q9CB91; GB:AL450380; NID:G13093927; PIDN:CA631818.1; GSPDB:C

C/Genetics:
 A/Gene: ML2302

C/Superfamily: regulatory protein fnr; cAMP receptor protein cyclic nucleotide-binding

Alignment Scores:
 Pred. No.: 126 Length: 224
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.29% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x B87197 (1-224)

QY 385 CAACCTCTGCGCGTGGCGCCGC 408
 Db 122 GlnLeuLeuArgValLeuAlaArg 129

RESULT 22

E70790

probable transcription regulator RV3676 - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: E70790
 R/Col: S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: E70790
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
 A/Residues: 1-224 <COL>

A/Cross-references: UNIPROT:O63644; GB:AL022121; GB:AL123456; NID:G3261559; PIDN:CAA1799

A/Experimental source: strain H37RV
 C/Genetics:

A/Gene: RV3676
 C/Superfamily: regulatory protein fnr; cAMP receptor protein cyclic nucleotide-binding

Alignment Scores:
 Pred. No.: 126 Length: 224
 Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.29% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x E70790 (1-224)

QY 385 CAACCTCTGCGCGTGGCGCCGC 408
 Db 122 GlnLeuLeuArgValLeuAlaArg 129

RESULT 23

T46967

diheme cytochrome soxB precursor [similarity] - Paracoccus denitrificans

N/Alternate names: cytochrome monoheme c-type
 C/Species: Paracoccus denitrificans

C/Date: 03-Nov-2000 #sequence_revision 03-Nov-2000 #text_change 09-Jul-2004

C/Accession: T46967
 R/Modara, C.; Bardischewsky, F.; Friedrich, C.G.

J. Bacteriol. 179, 5014-5023, 1997
 A/Title: Cloning and characterization of sulfite dehydrogenase, two c-type cytochromes, and sulfur oxidation.

A/Reference number: Z24324; MUID:97405897; PMID:9260941

A/Accession: T46967
 A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA
 A/Residues: 1-236 <MOD>

A/Cross-references: UNIPROT:O33435; EMBL:X79242; NID:G2253074; PIDN:CAA55828.1; PID:G2253

A/Experimental source: strain G817
 C/Genetics:

A/Gene: soxB
 C/Superfamily: Paracoccus denitrificans diheme cytochrome soxB; cytochrome c homology; C/

C/Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein
 F/1-18/Domain: signal sequence #status predicted <Sig>

F/19-236/Product: c-type cytochrome soxB #status predicted <Mat>
 F/30-127/Domain: cytochrome c homology <Cys>

F/163-232/Domain: cytochrome c6 homology <CY6>
 F/40/43/Binding site: heme (Cys) (covalent) #status predicted

F/44/109/Binding site: heme (His, Met) (axial ligands) #status predicted
 F/172/175/Binding site: heme (Cys) (covalent) #status predicted

F/176/213/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Alignment Scores:
 Pred. No.: 125 Length: 236
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.29% Indels: 0
 DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x T46967 (1-236)

QY 372 GCGGCTGCTGGCGCAACTCTGCG 395
 Db 14 AlaAlaAlaGlyAlaThrProAla 21

RESULT 24

PROB468

probable nuoD protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis
 C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: C70648
 R/Col: S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998
 A/Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: C70648
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA
 A/Residues: 1-262 <COL>

A/Cross-references: UNIPROT:O33435; EMBL:X79242; NID:G2253074; PIDN:CAA55828.1; PID:G2253

A/Experimental source: strain G817
 C/Genetics:

A/Gene: soxB
 C/Superfamily: Paracoccus denitrificans diheme cytochrome soxB; cytochrome c homology; C/

A/Cross-references: UNIPROT:P95172; GB:Z83867; GB:AL123456; NID:g3261695; PIDN:CAB06285.
A/Experimental source: strain H37Kv
C/Genetics:
A/Gene: nuoJ

Alignment Scores:

Pred. No.:	122	Length:	262
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x C70648 (1-262)

QY 494 AGAGGACAGAGGTAGCTGCGCTC 517

Db 103 ArgGlyGlnArgValAlaAlaVal 110

RESULT 25

AH0775

hydroxyethylthiazole kinase (EC 2.7.1.50) [imported] - *Salmonella enterica* subsp. *enterica*
C/Species: *Salmonella enterica* subsp. *enterica* serovar Typhi

A/Note: this species has also been called *Salmonella typhi*

C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C/Accession: AH0775

R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Conerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, P.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov

A/Reference number: AB0502; MUID:21534947; PMID:11677608

A/Accession: AH0775

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1265 <PAR>

A/Cross-references: GB:AL51382; PIDN:CAD02526.1; PID:g16503387; GSPDB:GN00176

C/Genetics:

A/Gene: STY2376

C/Superfamily: hydroxyethylthiazole kinase; hydroxyethylthiazole kinase homology

C/Keywords: phosphotransferase

Alignment Scores:

Pred. No.:	122	Length:	265
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.30%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x AH0775 (1-265)

QY 366 GCTCTGCGGCGACTGCGCGGGG 343

Db 150 AlaLeuAlaLeuProAlaAla 157

RESULT 26

AE0156

probable transcription regulatory protein YPO1279 [imported] - *Yersinia pestis* (strain C
C/Species: *Yersinia pestis*
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: AE0156

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AE0156

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-266 <KOR>

A/Cross-references: UNIPROT:Q8ZGU0; GB:AL590842; PIDN:CAC90112.1; PID:g15979332; GSPDB:GT
C/Genetics:
A/Gene: YPO1279
C/Superfamily: regulatory protein uxuR 2

Alignment Scores:

Pred. No.:	122	Length:	266
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x AE0156 (1-266)

QY 558 TCAGTGGAGACAGGCTCCCCC 581

Db 40 SerValGlyAspArgLeuProPro 47

RESULT 27

B70328

hypothetical protein aq_313 - *Aquifex aeolicus*
C/Species: *Aquifex aeolicus*

C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C/Accession: B70328

V. Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

Nature 392, 353-358, 1998

A/Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A/Reference number: A70300; MUID:98196666; PMID:9537320

A/Accession: B70328

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-279 <AOE>

A/Cross-references: UNIPROT:O6654; GB:AE000683; NID:g2982996; PIDN:AAC06617.1; PID:g298;

A/Experimental source: strain VFS

C/Genetics:

A/Gene: aq_313

C/Superfamily: hypothetical protein HP0152

Alignment Scores:

Pred. No.:	120	Length:	279
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x B70328 (1-279)

QY 988 GGCCTGGAGGCTGCTGCTGCTG 1011

Db 248 GlyArgGluAlaValArgLeuLeu 255

RESULT 28

T34848

probable transcription regulator - *Streptomyces coelicolor*
C/Species: *Streptomyces coelicolor*
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C/Accession: T34848

R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, February 1999

A/Reference number: Z21559

A/Accession: T34848

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-279 <OLI>

A/Cross-references: UNIPROT:Q9Z548; EMBL:AL035478; PIDN:CAB36602.1; GSPDB:GN00070; SCOEDE

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCOEDE:SC2G5.15c

Alignment Scores:

Pred. No.: 120 Length: 279
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x T34848 (1-279)

OY 802 GCGCGAGCGCTGGGCGCGCGG 779
 DB 162 AlalProAlaProGlyAlaAlaGly 169

RESULT 29

S46793
 vacuolar protein sorting protein - yeast (*Saccharomyces cerevisiae*)
 N/Alternate names: protein PEP1; protein VPS29; protein YHR012w
 C/Species: *Saccharomyces cerevisiae*
 C/Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
 C/Accession: S46793
 R/DU, 2.
 Submitted to the EMBL Data Library, June 1994
 A/Description: The sequence of *S. cerevisiae* cosmid L2825.
 A/Reference number: S46774
 A/Accession: S46793
 A/Molecule type: DNA
 A/Residues: 1-282 <DUZ>
 A/Cross-references: UNIPROT:P38759; EMBL:U10400; NID:G500701; PID:G50712; GSPDB:GN00008
 A/Gene: VPS29; MIPS:YHR012w
 A/Cross-references: MIPS:YHR012w; SGD:S0001054
 A/Map position: 8R
 A/Intons: 16/3
 C/Superfamily: VPS29-like phosphoesterase-related protein

Alignment Scores:
 Pred. No.: 120 Length: 282
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.29% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x S46793 (1-282)

OY 793 GCGCTGGCGGCGAGCTGGAAGG 816
 DB 120 AlalProAlaProGlyAlaAlaGly 127

RESULT 30

F48423
 homeotic protein engrailed 2 - chicken
 C/Species: *Gallus gallus* (chicken)
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Aug-2004
 C/Accession: F48423
 R/Logan, C.; Hanke, M.C.; Noble-Topham, S.; Nallathani, D.; Provart, N.J.; Joyner, A.L.
 Dev. Genet. 13, 345-358, 1992
 A/Title: Cloning and sequence comparison of the mouse, human, and chicken engrailed gene
 A/Reference number: A48423; MUID:93185339; PMID:1363401
 A/Accession: F48423
 A/Status: preliminary; not compared with conceptual translation
 A/Molecule type: DNA
 A/Residues: 1-288 <LOG>
 C/Superfamily: homeobox homology
 C/Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F/200-356/Domain: homeobox homology <HOX>

Alignment Scores:
 Pred. No.: 119 Length: 288
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x F48423 (1-288)

OY 808 GCTGCGCGCGGCGGCTGGGCGG 785
 DB 87 AlalAlaProAlaProGlyAla 94

RESULT 31

G87018
 probable membrane protein [imported] - *Mycobacterium leprae*
 C/Species: *Mycobacterium leprae*
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C/Accession: G87018
 R/Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Hoi
 R.; Davies, R.M.; Devlin, K.; Dutroy, S.; Feltwell, T.; Frazer, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
 A/Title: Massive gene decay in the leprosy bacillus.
 A/Reference number: A86909; MUID:21128732; PMID:11234002
 A/Accession: G87018
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-293 <STO>
 A/Cross-references: UNIPROT:Q9CCB9; GB:AL450380; NID:G13092950; PIDN:CA31258.1; GSPDB:G
 C/Genetics:
 A/Gene: mmpS3

Alignment Scores:
 Pred. No.: 119 Length: 293
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x G87018 (1-293)

OY 1078 CCCCTTCCTCTCCATCAGCAACA 1055
 DB 171 ProLeuProProProSerAlaThr 178

RESULT 32

C87667
 ABC transporter, ATP-binding protein CC3373 [imported] - *Caulobacter crescentus*
 C/Species: *Caulobacter crescentus*
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 16-Aug-2004
 C/Accession: C87667
 R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Taub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A/Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A/Reference number: A87249; MUID:21173698; PMID:11259647
 A/Accession: C87667
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-314 <STO>
 A/Cross-references: UNIPROT:Q9A333; GB:AE005673; NID:G13425079; PIDN:AAK25335.1; GSPDB:G
 C/Genetics:
 A/Gene: CC3373
 C/Superfamily: ATP-binding cassette homology

Alignment Scores:
 Pred. No.: 117 Length: 314
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x C87667 (1-314)

QY 363 CTCGCCGCACTGCGCGCGCTC 340
 |||||
 DB 244 LeuAlaAlaLeuProAlaAlaLeu 251

RESULT 33

S20880
 homeotic protein Hox 4.5 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 16-Sep-1992 #sequence revision 16-Sep-1992 #text_change 16-Aug-2004
 C:Accession: S20880; S09569; S09398
 R:Renucci, A.; Zappavigna, V.; Zakany, J.; Izpisua-Belmonte, J.C.; Buerki, K.; Duboule, EMO J. 11, 1459-1468, 1992
 A:Title: Comparison of mouse and human HOX-4 complexes defines conserved sequences invol
 A:Reference number: S20879; PMID:92224884; PMID:1348690
 A:Accession: S20880
 A:Molecule type: DNA
 A:Residues: 1-339 <REN>
 A:Cross-references: UNIPROT:P28357; EMBL:X62669; NID:951414; PIDN:CAA44542.1; PID:951416
 R:Duboule, D.; Dolle, P.
 EMO J. 8, 1497-1505, 1989
 A:Title: The structural and functional organization of the murine HOX gene family resen
 A:Reference number: S09569; PMID:89356621; PMID:2569969
 A:Accession: S09569
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Gene: Hox-4.5
 A:Introns: 260/1
 A:Superfamily: homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:273-329/Domain: homeobox homology <HOX>

Alignment Scores:

Pred. No.:	115	Length:	339
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x S20880 (1-339)

QY 481 TCACGCTCTCAAGAGGACAGAG 504
 |||||
 DB 173 SerSerSerSerLysArgThrGlu 180

RESULT 34

S18649
 homeotic protein HOX D9 - human
 N:Alternate names: homeotic protein Hox 4C; homeotic protein Hox 5.2
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 16-Aug-2004
 C:Accession: S18649; S05958; S14935; A32830
 R:Zappavigna, V.; Renucci, A.; Izpisua-Belmonte, J.C.; Utiel, G.; Peschle, C.; Duboule, EMO J. 10, 4177-4187, 1991
 A:Title: HOX4 genes encode transcription factors with potential auto- and cross-regulat
 A:Reference number: S18649; PMID:92097538; PMID:1756725
 A:Accession: S18649
 A:Molecule type: mRNA
 A:Residues: 1342 <ZAP>
 A:Cross-references: UNIPROT:P28356; EMBL:X59372; NID:932390; PIDN:CAA42016.1; PID:932391

A>Note: Intron position was determined by sequencing of genomic DNA
 R:Oliver, G.; Sidell, N.; Fiske, W.; Heinemann, C.; Mohandas, T.; Sparkes, R.S.; De Rober
 Genes Dev. 3, 641-650, 1989
 A:Title: Complementary homeo protein gradients in developing limb buds.
 A:Reference number: A32830; PMID:89306602; PMID:2569311
 A:Accession: S05958

A:Molecule type: DNA
 A:Residues: 264-265, 'A', 267-342 <OLI>
 A:Cross-references: EMBL:X15506; NID:932397; PIDN:CAA33528.1; PID:932398
 R:Campana, D.; d'Esposito, M.; Falella, A.; Pannese, M.; Mignolacci, E.; Morelli, F.; St
 Nucleic Acids Res. 17, 10385-10402, 1989
 A:Title: The human HOX gene family.
 A:Reference number: S07541; PMID:90098876; PMID:2574852
 A:Accession: S14935
 A:Molecule type: DNA
 A:Residues: 275-340 <ACA>
 C:Genetics:

A:Gene: GDB:HOXD9
 A:Cross-references: GDB:120678; OMIM:142982
 A:Map position: 2q31-2q31
 A:Introns: 263/1
 C:Superfamily: homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:276-332/Domain: homeobox homology <HOX>

Alignment Scores:

Pred. No.:	115	Length:	342
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x S18649 (1-342)

QY 481 TCACGCTCTCAAGAGGACAGAG 504
 |||||
 DB 175 SerSerSerSerLysArgThrGlu 182

RESULT 35

T08826
 secretory carrier membrane protein homolog propin - human
 C:Species: Homo sapiens (man)
 C:Date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text_change 09-Jul-2004
 C:Accession: T08826
 R:Winfield, S.L.; Tayebi, N.; Martin, B.M.; Gims, E.I.; Sidransky, E.
 Genome Res. 7, 1020-1026, 1997
 A:Title: Identification of three additional genes contiguous to the glucocerebrosidase 1c
 A:Reference number: Z1482; PMID:97474796; PMID:9331372
 A:Accession: T08826
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-347 <WIN>
 A:Cross-references: UNIPROT:O14828; EMBL:AF023268; NID:92564910; PIDN:AC51821.1; PID:925

A:Gene: Propin
 A:Map position: 1
 A:Introns: 22/3; 48/3; 89/3; 130/1; 173/1; 226/2; 260/2; 299/3

Alignment Scores:

Pred. No.:	115	Length:	347
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.30%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x T08826 (1-347)

QY 1081 GCCCCCTCTCTCCATCAGCA 1058
 |||||
 DB 59 AlaProLeuProProSerAla 66

RESULT 36

T34931
 hypothetical protein SC3F9.09 SC3F9.09 - Streptomyces coelicolor
 C/Species: Streptomyces coelicolor
 C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C/Accession: T34931
 R/Seeger, K.U.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1998
 A/Reference number: Z21562
 A/Accession: T34931
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-363 <SEE>
 A/Cross-references: UNIPROT:O69949; EMBL:AL023862; PIDN:CA19632.1; GSPDB:GN00070; SCOPED
 A/Experimental source: strain A3(2)
 C/Genetics:
 A/Gene: SC0DB:SC3F9.09

Alignment Scores:

Pred. No.:	113	Length:	363
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x T34931 (1-363)

QY 130 CTATCCGGCTGACCCCGGCCCG 153

Db 269 LeuSerLysThrProAlaPro 276

RESULT 37

T51339
 mitogen-activated protein kinase kinase (EC 2.7.1.1-) 4 [validated] - Arabidopsis thaliana
 N/Alternate names: MAP kinase kinase 4
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C/Accession: T51339
 R/Chimura, K.; Mizoguchi, T.; Hayashida, N.; Seki, M.; Shinzaki, K.
 DNA Res. 5, 341-348, 1998
 A/Title: Molecular cloning and characterization of three cDNAs encoding putative mitogen
 A/Reference number: Z25272; MUID:99156228; PMID:10048483
 A/Accession: T51339
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-366 <ICH>
 A/Cross-references: UNIPROT:O80397; EMBL:AB015315; PIDN:BA28830.1
 A/Experimental source: strain Columbia
 C/Genetics:
 A/Gene: ATMKK4
 C/Function:
 A/Description: (EC 2.7.1.1-); mitogen-activated protein kinase kinase [validated, MUID:99
 C/Superfamily: kinase-related transforming protein; protein kinase homology
 C/Keywords: phosphotransferase; protein kinase

Alignment Scores:

Pred. No.:	113	Length:	366
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.30%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x T51339 (1-366)

QY 919 AGTCGCCGAGAGGCTCCAGAT 896

Db 16 SerArgProArgArgArgProArg 23

RESULT 38

AE2227
 transposase alr3172 [imported] - Nostoc sp. (strain PCC 7120)

C/Species: Nostoc sp. PCC 7120

A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C/Accession: AE2227
 R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AE2227
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-374 <RUR>
 A/Cross-references: UNIPROT:Q8YR51; GB:BA000019; PIDN:BA075071.1; PID:G17132467; GSPDB:G
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: alr3372

Alignment Scores:

Pred. No.:	113	Length:	374
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x AE2227 (1-374)

QY 1798 MAGGAAATGCAAGGCTTTGGGGG 1821

Db 270 LysGlyMetGlnArgLeuTrpGly 277

RESULT 39

T05707
 phosphate transport protein G7, mitochondrial - soybean
 C/Species: Glycine max (soybean)
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C/Accession: T05707
 R/Takabatake, R.; Hata, S.; Taniguchi, M.; Kouchi, H.; Sugiyama, T.; Izui, K.
 submitted to the EMBL Data Library, July 1998
 A/Description: Isolation and characterization of cDNAs encoding mitochondrial phosphate t
 A/Reference number: Z15423
 A/Accession: T05707
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-375 <YAK>
 A/Cross-references: UNIPROT:O80412; EMBL:AB016063; PIDN:BA031562.1
 A/Experimental source: cultivar Akisengoku; root tip
 C/Genetics:
 A/Genome: nuclear
 C/Superfamily: ADP,ATP carrier protein; ADP,ATP carrier protein repeat homology
 C/Keywords: mitochondrion; phosphate transport

Alignment Scores:

Pred. No.:	113	Length:	375
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.30%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x T05707 (1-375)

QY 811 CCAGCTGCCGCGCAGCGCTGGG 788

Db 362 ProAlaAlaAlaProAlaProGly 369

RESULT 40

C75580
 adenine deaminase-related protein - Deinococcus radiodurans (strain R1)
 C/Species: Deinococcus radiodurans
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: C75580

R.White, O.; Eiseen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathavan, J.U.; Lan, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; NCID:20036896; PMID:10567266
A:Accession: C75580
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-376 <WHI>
A:Cross-references: UNIPROT:Q9RYP2; GB:AE001863; GB:AE001825; NCID:96460670; PIDN:AAF1237
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0268
A:Map position: 2

Alignment Scores:
Pred. No.: 113 Length: 376
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x C75580 (1-376)

QY 930 CGCCCTGCTGCAGCCCTGGCGG 953
Db 65 ArgProAlaAlaAlaProAlaGly 72
|||||

RESULT 41
S08296
beta-lactamase (EC 3.5.2.6) precursor - Citrobacter freundii (strain GN346)
N:Alternate names: cephalosporinase
C:Species: Citrobacter freundii
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: S08296; A60909; A27028; A24869
R:Tsukamoto, K.; Tachibana, K.; Yamazaki, N.; Ishii, Y.; Ujite, K.; Nishida, N.; Sawai,
Eur. J. Biochem. 188, 15-22, 1990
A>Title: Role of lysine-67 in the active site of class C beta-lactamase from Citrobacter
A:Reference number: S08296; NCID:90201023; PMID:1969344
A:Accession: S08296
A:Molecule type: DNA
A:Residues: 1-381 <TSU>
A:Cross-references: UNIPROT:P05193; GB:X51632; NCID:940457; PIDN:CAA35959.1; PID:G40458
A>Note: part of this sequence was confirmed by amino acid sequencing
R:Sawai, T.; Yamaguchi, A.; Tsukamoto, K.
Rev. Infect. Dis. 10, 721-725, 1988
A>Title: Amino acid sequence, active-site residue, and effect of suicide inhibitors on C
A:Reference number: A60909; NCID:89043488; PMID:3263684
A:Accession: A60909
A>Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 21-97, 'T', 99-381 <SAW>
R:Yamaguchi, A.; Adachi, H.; Sawai, T.
FEBS Lett. 218, 126-130, 1987
A>Title: Identification of the active site of Citrobacter freundii beta-lactamase using
A:Reference number: A27028; NCID:87247241; PMID:3496243
A:Accession: A27028
A:Molecule type: protein
A:Residues: 66-87 <YAM>
R:Uindberg, F.; Normark, S.
Eur. J. Biochem. 156, 441-445, 1986
A>Title: Sequence of the Citrobacter freundii O560 chromosomal ampc beta-lactamase gene.
A:Reference number: A24869; NCID:8612473; PMID:3486121
A:Accession: A24869
A:Molecule type: DNA
A:Residues: 1-96, 'R', 98-142, 'G', 144, 'V', 146-149, 'E', 151-184, 'S', 186-223, 'L', 225-242, 'V',
A:Cross-references: GB:X03866; NCID:940451; PIDN:CAA27494.1; PID:G40452
A:Experimental source: wild-type isolate O560
C:Genetics:
A:Gene: ampc
C:Superfamily: Escherichia coli beta-lactamase

C:Keywords: antibiotic resistance; hydrolase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-381/Product: beta-lactamase #status predicted <MAT>
F:84,87/Active site: Ser, Lys #status experimental

Alignment Scores:
Pred. No.: 112 Length: 381
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x S08296 (1-381)

QY 369 GTTGCTGCGCCGCACTGCGCGG 346
Db 311 ValAlaLeuAlaAlaLeuProAla 318
|||||

RESULT 42
S45109
beta-lactamase (EC 3.5.2.6) precursor - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S45109
R:Tzouvelektis, L.S.; Tzelepi, E.; Mentis, A.F.
submitted to the EMBL Data Library, March 1994
A:Description: Nucleotide sequence of a plasmidic cephalosporinase gene (bla-LAT-1) found
A:Reference number: S45109
A:Accession: S45109
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-381 <TZO>
A:Cross-references: UNIPROT:Q48443; EMBL:X78117; NCID:9496632; PIDN:CAA5007.1; PID:G49663
C:Superfamily: Escherichia coli beta-lactamase
C:Keywords: hydrolase
F:84/Active site: Ser #status predicted

Alignment Scores:
Pred. No.: 112 Length: 381
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x S45109 (1-381)

QY 369 GTTGCTGCGCCGCACTGCGCGG 346
Db 311 ValAlaLeuAlaAlaLeuProAla 318
|||||

RESULT 43
AD0897
conserved hypothetical protein STY3429 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typh
A>Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0897
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; NCID:21534947; PMID:11677608
A:Accession: AD0897
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-381 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07770.1; PID:G16504319; GSPDB:GN00176
C:Genetics:
A:Gene: STY3429

C/Superfamily: yhad protein

Alignment Scores:

Pred. No.: 112 Length: 381
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x AD0897 (1-381)

QY 783 GCGGCCCGAGCGCTGCGCGCA 806

DB 243 AAlaAlaProGlyAlaGlyAlaAla 250

RESULT 44

protein F47D12.1 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C/Accession: D88479

R/Anonymous, The C. elegans Sequencing Consortium.

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

Science 282, 2012-2018, 1998

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see webstes genome.wustl.edu/geno/C_elegans/ and www.sanger.ac.uk/Projects/C_ele

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A/Accession: D88479

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-387 <STO>

A/Cross-references: UNIPROT:Q8WQ41; GB:chr_III; P1DN:AA64313.1; PID:g272253; GSPDB:GN00

C/Genetics:

A/Gene: F47D12.1

A/Map position: 3

Alignment Scores:

Pred. No.: 112 Length: 387
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x D88479 (1-387)

QY 1670 ACAGCAGGGGGAAGTGAGGAGA 1647

DB 213 ThrAlaGlyGlyValAlaArg 220

RESULT 45

hypothetical protein CC3439 [imported] - Caulobacter crescentus

C/Species: Caulobacter crescentus

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: E87675

R/Hierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A/Title: Complete Genome Sequence of Caulobacter crescentus.

A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: E87675

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-394 <STO>

A/Cross-references: UNIPROT:Q9A2M7; GB:AE005673; NID:G1342515; P1DN:AAK25401.1; GSPDB:G

C/Genetics:

A/Gene: CC3439

Alignment Scores:

Pred. No.: 111 Length: 394

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x E87675 (1-394)

QY 808 GGTGCCCGCAGCGCTGAGGCC 785

DB 13 AAlaAlaProAlaProAlaAla 20

RESULT 46

T08345 hypothetical protein H1502 [imported] - Halobacterium sp. (strain NRC-1) plasmid pNRC100

C/Species: Halobacterium sp.

A/Variety: strain NRC-1

C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C/Accession: T08345

R/Mg, W.V.; Clufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, B.; I

Genome Res. 8, 1131-1141, 1998.

A/Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or m

A/Reference number: Z16406; MUID:99063795; PMID:9847077

A/Status: translated from GB/EMBL/DBJ

A/Accession: T08345

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-397 <DAS>

A/Cross-references: UNIPROT:O52032; EMBL:AF016485; NID:g2822276; PID:g2822406; HALOSP:H1

A/Experimental source: strain NRC-1

C/Genetics:

A/Gene: HALOSP:H1502

A/Genome: plasmid pNRC100

C/Superfamily: Halobacterium plasmid pNRC100 hypothetical protein H1502

Alignment Scores:

Pred. No.: 111 Length: 397
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x T08345 (1-397)

QY 753 GCCAGCCTTGAGCAGCGCTG93C 776

DB 85 AAlaSerLeuGlyAlaGlyArgGly 92

RESULT 47

SYZMW1 naringenin-chalcone synthase (EC 2.3.1.74) whp1 - maize

C/Species: Zea mays (maize)

C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C/Accession: S16599

R/Franken, P.; Niesbach-Kloeegen, U.; Weydemann, U.; Marechal-Drouard, L.; Saedler, H.; V

EMBO J. 10, 2605-2612, 1991

A/Title: The duplicated chalcone synthase genes C2 and whp (white pollen) of Zea mays are

gene in.

A/Reference number: S16598; MUID:91330885; PMID:1714383

A/Accession: S16599

A/Molecule type: DNA

A/Residues: 1-400 <FRA>

A/Cross-references: UNIPROT:P24824; EMBL:X60204; NID:g22511; P1DN:CAA42763.1; PID:g22512

C/Genetics:

A/Gene: whp1

A/Map position: 2L

A/Intons: 64/1

C/Superfamily: Type III polyketide synthase

C/Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Alignment Scores:

Pred. No.: 111 Length: 400
Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.30% Indels: 0
 DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x SYZMM1 (1-400)

QY 427 GGTGGCGGACGAGGTGTCGGCGG 404
 DB 221 GYLAALAAIAGLYArgGlygly 228

RESULT 48

A45056

osteogenic protein 2 precursor - human

N.Alternate names: bone morphogenetic protein 8; OP-2

C.Species: Homo sapiens (man)

C.Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C.Accession: A45056

R.Oezkaynak, E.; Schmeigelberg, P.N.; Jin, D.F.; Clifford, G.M.; Warren, F.D.; Drier, E.

J. Biol. Chem. 267, 25220-25227, 1992

A.Title: Osteogenic protein-2. A new member of the transforming growth factor-beta super

A.Reference number: A45056; MUID:93094231; PMID:1460021

A.Accession: A45056

A.Status: preliminary; not compared with conceptual translation

A.Molecule type: nucleic acid

A.Residues: 1-402 <OZK>

A.Cross-references: UNIPROT:P34820; GB:M97016; NID:g189389; PIDN:AAB01360.1; PID:g189390

A.Experimental source: hippocampus

A.Note: sequence extracted from NCBI backbone (NCBIP.120189)

C.Genetics:

A.Gene: GDB:BMP8; OP-2

A.Cross-references: GDB:136392

C.Superfamily: Inhibin

Alignment Scores:

Pred. No.:	111	Length:	402
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x A45056 (1-402)

QY 134 CCGGGTCGACCCGCGCCGCTGCT 157
 DB 55 ProGLYArgProArgProArgala 62

RESULT 49

H69812

conserved hypothetical protein yfmi - Bacillus subtilis

C.Species: Bacillus subtilis

C.Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C.Accession: H69812

R.Kunert, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beretx

C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Eutian, K.D.; Birlington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A.Authors: Foulger, D.; Henatz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

iech, J.; Harwood, C.R.; Henatz, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F.

Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maues

Y, M.; Ogawa, K.; Ogilwara, A.; Odega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A.; Schleich, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tempstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K

A.Authors: Yoshikawa, H.F.; Zumein, E.; Yoshikawa, H.; Danchin, A.

A.Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A.Reference number: A69580; MUID:98044033; PMID:9384377

A.Accession: H69812

A.Status: preliminary; nucleic acid sequence not shown; translation not shown

A.Residues: 1-406 <KUN>
 A.Cross-references: UNIPROT:O34440; GB:Z99108; GB:AL009126; NID:g2633055; PIDN:CAB12575.1
 A.Experimental source: strain 168
 C.Genetics:
 A.Gene: yfmi

Alignment Scores:

Pred. No.:	111	Length:	406
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x H69812 (1-406)

QY 1518 ACCAACAACCACTGGCTTCAGC 1541
 DB 13 ThrThrThrAsnLeuGlyPheSer 20

RESULT 50

E70475

folylpolyglutamate synthetase - Aquifex aeolicus

C.Species: Aquifex aeolicus

C.Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C.Accession: E70475

R.Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove

Nature 392, 353-358, 1998

A.Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A.Reference number: A70300; MUID:98196666; PMID:9537320

A.Accession: E70475

A.Status: preliminary; nucleic acid sequence not shown; translation not shown

A.Molecule type: DNA

A.Residues: 1-410 <AQF>

A.Cross-references: UNIPROT:O67833; GB:AE000769; NID:g2984262; PIDN:AAC07789.1; PID:g2984

A.Experimental source: strain VFS

C.Genetics:

A.Gene: folC

C.Superfamily: folylpolyglutamate synthase

Alignment Scores:

Pred. No.:	110	Length:	410
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.30%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x E70475 (1-410)

QY 1510 GAGGTCTGAGAGAGAGCCCTTA 1487
 DB 284 GluValLeuArgGluLysProLeu 291

Search completed: March 22, 2005, 03:04:20
 Job time : 91.1055 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 02:34:17 ; Search time 327.771 Seconds
(without alignments)
5883.642 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 620
Sequence: 1 aggcgcataacatagagaag.....ctcgaaagggcataggtc999 1883

Scoring table:
Oligo
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 segs, 512079187 residues

Word size: 1

Total number of hits satisfying chosen parameters: 322408

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 100 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=x1p
-Q/cgnt2_1/USFTO.spool_p/US10030271/runat_21032005.153546.26822/app_query.fasta_1.3150
-DB=Uniprot_03 -QFWT=faetan -SUFFIX=oligo.rup -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10030271 @CNC 1.1.578 @runat 21032005.153546.26822 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Uniprot 03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	247	39.8	326	1 DED2_HUMAN	08wz8 homo sapien
2	135	21.8	330	1 DED2_MOUSE	08qzv0 mus musculu
3	23	3.7	168	2 Q8BRM9	08brn9 mus musculu
4	14	2.3	369	2 Q919M3	0919m3 brachydanio
5	13	2.1	318	1 DEDD_HUMAN	075618 homo sapien
6	13	2.1	318	1 DEDD_MOUSE	09z113 mus musculu
7	13	2.1	318	1 DEDD_RAT	09z260 ratcus norv
8	13	2.1	404	2 Q6DHN2	06dhn2 brachydanio
9	11	1.8	243	2 Q6GNZ8	06gnz8 xenopus lae
10	10	1.6	218	2 Q91M13	091m13 molluscum c
11	10	1.6	235	2 Q98222	098222 molluscum c
12	10	1.6	244	2 Q6DHV2	06dhv2 brachydanio
13	10	1.6	581	2 Q9F1E2	09f1e2 acetabacter
14	10	1.6	729	2 Q91TW1	091tw1 tupaid her
15	10	1.6	1461	1 I818_PRIVIF	P1675 pseudorabie
16	10	1.6	3247	2 Q65553	065553 bovine herp

17	10	1.6	3247	2 Q77CD4	077cd4 bovine herp
18	9	1.5	110	2 Q8W615	08w615 sinorhizobi
19	9	1.5	115	2 Q8C518	08c518 mus musculu
20	9	1.5	127	2 Q9Y9Y9	09y9y9 aeropyrum p
21	9	1.5	148	2 Q67YU8	067yu8 oryza sativ
22	9	1.5	201	2 Q6ZLH4	06zlh4 oryza sativ
23	9	1.5	222	2 Q671V3	0671v3 oryza sativ
24	9	1.5	227	2 Q706G3	0706g3 anopheles g
25	9	1.5	241	2 Q6MBX0	06mbx0 paracalamyd
26	9	1.5	243	2 Q6PML5	09f15 oryza sativ
27	9	1.5	274	2 Q73V06	073v06 mycobacteri
28	9	1.5	284	2 Q635Y8	0635y8 burkholderi
29	9	1.5	287	2 Q855F3	0855f3 mycobacteri
30	9	1.5	296	2 Q6YTS7	06yts7 oryza sativ
31	9	1.5	307	2 Q9RYL0	09ryl0 deinococcus
32	9	1.5	310	2 Q84Z99	084z99 oryza sativ
33	9	1.5	335	2 Q89402	089402 bradyrhizob
34	9	1.5	347	2 Q8R353	08r353 mus musculu
35	9	1.5	370	1 WNT1_HUMAN	08w628 homo sapien
36	9	1.5	373	2 Q7Z5U1	07z5u1 homo sapien
37	9	1.5	374	2 Q6PDP7	06pdp7 homo sapien
38	9	1.5	375	2 Q6RYV4	06ryv4 homo sapien
39	9	1.5	382	2 Q82B67	082b67 streptomyce
40	9	1.5	385	2 Q6M204	06m204 coxynebacte
41	9	1.5	393	2 Q7XJW1	07xjw1 oryza sativ
42	9	1.5	395	2 Q8NL75	08nl75 coxynebacte
43	9	1.5	410	2 Q7MMW4	07mmw4 bordetella
44	9	1.5	447	2 Q7Q5N2	07q5n2 anopheles g
45	9	1.5	452	2 Q7VUX6	07vux6 bordetella
46	9	1.5	454	2 Q8YXAX	08yxax anabaena sp
47	9	1.5	458	2 Q8BV76	08bv76 mus musculu
48	9	1.5	465	1 FUMC_SYNPX	0910p9 streptomyce
49	9	1.5	468	2 Q9L0P9	09l0p9 streptomyce
50	9	1.5	478	2 Q82HT1	082ht1 xenopus lae
51	9	1.5	516	2 Q6DDP9	06ddp9 xenopus lae
52	9	1.5	525	2 Q6K663	06k663 oryza sativ
53	9	1.5	571	2 Q75EB6	075eb6 ashya goss
54	9	1.5	614	1 GIDA_UREPA	09pr6 ureaplasma
55	9	1.5	629	2 Q871T5	0871t5 neosporea
56	9	1.5	636	2 Q6XPZ5	06xpz5 triticum ae
57	9	1.5	645	2 Q9D620	09d620 mus musculu
58	9	1.5	675	2 Q89YB1	089yb1 bradyrhizob
59	9	1.5	709	2 Q8BZN7	08bzn7 mus musculu
60	9	1.5	955	1 T150_HUMAN	09y2w1 homo sapien
61	9	1.5	1099	2 Q7XW40	07xw40 oryza sativ
62	9	1.5	1176	2 Q8PF66	08pf66 xanthomonas
63	9	1.5	1725	2 Q8UW61	08uw61 oryzae lat
64	9	1.5	5068	2 Q63LK7	063lk7 burkholderi
65	9	1.5	5835	2 Q63LK8	063lk8 burkholderi
66	8	1.3	27	2 Q6TQ89	06tq89 hepatitis c
67	8	1.3	27	2 Q6W5U5	06w5u5 hepatitis c
68	8	1.3	61	2 Q6KZ28	06kz28 oryza sativ
69	8	1.3	63	1 COM3_CONTE	09nda6 consus texti
70	8	1.3	63	2 Q931Y4	0931y4 streptomyce
71	8	1.3	68	2 Q75HL5	075hl5 oryza sativ
72	8	1.3	72	2 Q8HTZ8	08htz8 oryza sativ
73	8	1.3	75	2 Q8UH77	08uh77 agrobacteri
74	8	1.3	79	2 Q8LH71	08lh71 oryza sativ
75	8	1.3	81	2 Q6FE11	06fe11 oryza sativ
76	8	1.3	82	2 Q656Q0	0656q0 oryza sativ
77	8	1.3	83	2 Q96W43	096w43 ophiostoma
78	8	1.3	86	1 EXOX_RHILP	P14801 rhizobium 1
79	8	1.3	86	2 Q6W4V6	06w4v6 patriciella
80	8	1.3	89	2 Q9R9X5	09r9x5 pseudomonas
81	8	1.3	89	2 Q884W3	0884w3 pseudomonas
82	8	1.3	89	2 Q88E09	088e09 pseudomonas
83	8	1.3	90	2 Q7NTQ0	07ntq0 chromobacte
84	8	1.3	91	2 Q9N243	09n243 streptococ
85	8	1.3	91	2 Q9N244	09n244 nasalis lar
86	8	1.3	91	2 Q9N245	09n245 macaca mula
87	8	1.3	91	2 Q9N246	09n246 papio hamad
88	8	1.3	91	2 Q6ZLN8	06zln8 oryza sativ
89	8	1.3	93	2 Q9N5C5	09n5c5 caenorhabdi

90	8	1.3	93	2	Q9P1P2	Q9f1f2 enterococcu
c 91	8	1.3	96	2	Q6EUB0	Q6eub0 oryza sativ
92	8	1.3	98	2	Q9X6R6	Q9x6r6 micromonosp
c 93	8	1.3	98	2	Q89XY5	Q89xy5 bradyrhizob
94	8	1.3	100	2	Q9UZF4	Q9ufz4 caenorhabdi
c 95	8	1.3	101	2	Q6UPP9	Q6upp9 human immu
c 96	8	1.3	104	2	Q69KA4	Q69ka4 oryza sativ
97	8	1.3	105	2	Q92KN5	Q92kn5 rhizobium m
c 98	8	1.3	106	2	Q66VG6	Q66vg6 homo sapien
c 99	8	1.3	107	2	Q9LEI7	Q9lei7 hordeum vul
c 100	8	1.3	108	2	Q8S5E3	Q8sse3 oryza sativ

ALIGNMENTS

RESULT 1
DE2D2_HUMAN STANDARD: PRT: 326 AA.
Q8WXF8; Q8NBR2; Q8NBR2; Q8TA8; Q96D35;
05-JUL-2004 (Rel. 44, Created)
05-JUL-2004 (Rel. 44, Last sequence update)
05-JUL-2004 (Rel. 44, Last annotation update)
DNA-binding death effector domain-containing protein 2 (DED-containing protein FLAME-3).
Name=DED2; Synonym=FLAME3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, AND INTERACTION WITH CASP8.
MEDLINE=21850646; PubMed=11741985; DOI=10.1074/jbc.M110749200;
Roth W., Steiner-Lewen F., Pawlowski K., Godzik A., Reed J.C.;
RT "Identification and characterization of DED2, a death effector domain-containing protein."
J. Biol. Chem. 277:7501-7508(2002).
[2]
SEQUENCE FROM N.A. (ISOFORM 1), TISSUE SPECIFICITY, SUBCELLULAR LOCATION, AND INTERACTIONS WITH CASP8 AND GTF3C3.
MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd/4401038;
Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T., Alnemri E.S.;
RT "Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the FflIIC102 subunit of human transcription factor IIC1."
Cell Death Differ. 9:439-447(2002).
[3]
SEQUENCE FROM N.A. (ISOFORM 2).
MEDLINE=2220288; PubMed=1225123; DOI=10.1083/jcb.200112124;
Lee J.C., Schickling O., Stegh A.H., Oshima R.G., Dinsdale D., Cohen G.M., Peter M.E.;
RT "DED regulates degradation of intermediate filaments during apoptosis."
J. Cell Biol. 158:1051-1066(2002).
[4]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Teratocarcinoma;
PubMed=14702039; DOI=10.1038/ng1285;
Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Maeda H., Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagatani K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Puruya T., Kikawa E., Omuura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Niimiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hirotaka S., Chiba Y., Iehara S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O., Nomura Y., Togayari K., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,

Yoshikawa Y., Matsumawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Komiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wabe H., Hisigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukushima Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y., Okamoto S., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togoishi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human cDNAs."
Nat. Genet. 36:40-45(2004).
[5]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
TISSUE=Brain, and Lung;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Pangue C., Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullen S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.W., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez A., Rodriguez R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalske U., Smalls D.E., Scherren A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
INTERACTIONS WITH CASP8 AND CASP10.
PubMed=12527898; DOI=10.1038/sj.onc.1206099;
Alcivar A., Hu S., Tang Y., Yang X.;
RT "DED and DEDD associate with caspase-8/10 and signal cell death."
Oncogene 22:291-297(2003).
[7]
FUNCTION: May play a critical role in death receptor-induced apoptosis and may target CASP8 and CASP10 to the nucleus. May regulate degradation of intermediate filaments during apoptosis. May play a role in the general transcription machinery in the nucleus and might be an important regulator of the activity of GTF3C3.
SUBUNIT: Interacts with CASP8, CASP10 and GTF3C3. Homodimerizes and heterodimerizes with DEDD.
SUBCELLULAR LOCATION: Nuclear, accumulated in subnuclear structures resembling nucleoli.
[8]
ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1;
IsoId=Q8WXF8-1; Sequence=Displayed;
Name=2;
IsoId=Q8WXF8-2; Sequence=VSP_010312;
[9]
TISSUE SPECIFICITY: Expressed in most tissues. High levels were found in liver, kidney, heart, ovary, spleen, testes, skeletal muscle and peripheral blood leukocytes. Expression is absent or low in colon and small intestine. Expression is relatively high in the tumor cell lines chronic myelogenous leukemia K-562 and the colorectal adenocarcinoma SW480. Expression is moderate in the cervical carcinoma HeLa, the Burkitt's lymphoma Raji, the lung carcinoma A549, and the melanoma G361. In contrast, two leukemia cell lines, HL-60 (promyelocytic leukemia) and MOLT-4 (lymphoblastic leukemia), show relatively low levels.

```
CC -I DOMAIN: Interactions with Cas9 and CasP10 are mediated by the DED domain.
```

```
CC -I SIMILARITY: Contains 1 death effector (DED) domain.
```

```
CC -I CAUTION: Ref.5 (AAH13372) sequence differs from that shown due to a frameshift in position 186.
```

```
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/, or send an email to license@isb-sib.ch).
```

```
DR EMBL; AF443591; AAL48220.1; -
```

```
DR EMBL; AF457575; AAM10835.1; -
```

```
DR EMBL; AY125488; AAM95240.1; -
```

```
DR EMBL; AK075328; BAC11551.1; -
```

```
DR EMBL; BC023372; AAH13372.2; ALT_FRAME.
```

```
DR EMBL; BC027930; AAH27930.1; --
```

```
DR Genew; HGNC:24450; DED2.
```

```
DR H-invDB; HIX0015171; -
```

```
DR InterPro; IPR011029; DEATH_like.
```

```
DR InterPro; IPR001875; DED.
```

```
DR Pfam; PF01335; DED; 1.
```

```
DR SMART; SMO0031; DED; 1.
```

```
DR PROSITE; PS50168; DED; 1.
```

```
KW Alternative splicing; Apoptosis; DNA-binding; Nuclear protein; Transcription regulation.
```

```
FT DOMAIN 25 104
```

```
FT FT 104 109
```

```
FT DOMAIN 155 173
```

```
FT FT 145 149
```

```
FT VASPLIC 145 149
```

```
FT CONFLICT 27 27
```

```
FT CONFLICT 56 56
```

```
FT CONFLICT 79 79
```

```
FT CONFLICT 207 207
```

```
FT CONFLICT 230 230
```

```
SQ SEQUENCE 326 AA; 36178 MW; 3F7B0B307C870CD CRC64;
```

```
Alignment Scores:
```

```
Pred. No.: 1.34e-230 Length: 326
```

```
Score: 247.00 Matches: 247
```

```
Percent Similarity: 100.00% Conservative: 0
```

```
Best local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 39.84% Indels: 0
```

```
Gaps: 0
```

```
US-10-030-271-3 (1-1863) x DED2_HUMAN (1-326)
```

```
OY 361 GAGAGCAACTGCGGCTGTGGGGAACCTCCTGCGCGTGCTGCCGCCGACGACTGCTG 420
```

```
Dd |||||
```

```
80 GlnSerAsnLeuAlaArgLeuLeuSylGlnLeuLeuValLeuAlaArgHisAspLeuLeu 99
```

```
OY 421 CCGCACCTGGCGCGGCAAGCGCGCGCGGCACAGTGTCTCCAGAACGCTATAGCTATGGACAC 480
```

```
Dd |||||
```

```
100 ProHisLeuAlaArgLyAlaArgArgrProValSerProGluAlaArgTySerLygLythr 119
```

```
OY 481 TTCAGCTTTCAAAGAAGACAAGAGGTAGTCCTCCGTCGCGGAGCATGAACGACTTCT 540
```

```
Dd |||||
```

```
120 SerSerSerSerLyAlaArgThrGlnIleLySerCybaArgArgrArgIleSerSerSer 139
```

```
OY 541 GCAATTCTTCAGCAGAGGTCAATGGAGACAGGCTCCCCCAACAAGAGCGGAGGCGGG 600
```

```
Dd |||||
```

```
140 AlaAsnSerGlnGlnGlyGlnTrpLeuThrLySerProProThrLyArgGlnArgArg 159
```

```
OY 601 AGTCGGGCGCGGCGCACGATGTGTGTCCACGCGCGGAGAGAGGAGCGGCACCGCACCC 660
```

```
Dd |||||
```

```
160 SerArgGlyArgProSerGlyGlyAlaArgArgrArgArgrGlyAlaProAlaArgPro 179
```

```
OY 661 CAGCAGCAGTCAAGACCGCGCACGACTTCTCTGAAGCAAGTAAGTACCTGTGACATCCGG 720
```

Db	180	GlnInGlnInserGIuPProAlaArgProSerSerClnGluLysValThrCySaPILleArg	199
QY	721	CTCCGGGTTCCAGCAGAGTACTCGAGACATCGGACCTTTGAGCAGGGCGTGGCATCC	780
Db	200	LeuArgValArgAlaGluArgLysCysGluHISglYProAlaLeuGlnGlnGlnValAlaSer	219
QY	781	CGGGGGCCCCAGGCGCTGGCGGGCAGCTGGAGCTGTTTGGGCGAGGCCACCGCAGTGTCTG	840
Db	220	ArgArgProGlnAlaLeuAlaValArgGlnLeuAspAlaPheGlyGlnAlaThrAlaValLeu	239
QY	841	CGCTCAAGGGGACCTCGGGCTGTGTGGTTTGGACATCAAGTTCTCAGAGCTCTCTTACTGT	900
Db	240	ArgSerArgAspLeuGlySerValValCysAspIleLysPheSerGluLeuSerTyrLeu	259
QY	901	GACGCGCTTCTGGGGCGACTACCTAGTGGCGCCCTGTGTGACAGGCCCTGGGGCGGTTC	960
Db	260	AspAlaPheTrpGlyAspTyrLeuSerGlyValAlaLeuGlnAlaLeuArgGlyValPhe	279
QY	961	CTGACTGAGGCGCTTCGCGAGAGGCTGTGGCGCGGAGGCTGTTGCGCTGCTGTGATGTGTG	1020
Db	280	LeuThrGlnAlaLeuArgGlnAlaValAlaGlyValArgGluAlaValArgLeuLeuValSerVal	299
QY	1021	GATAGGCTGACTATGAGGCGTGGCGGGCGCGCGCTGTGTGCATGTGAGAGGAGGAGGGGG	1080
Db	300	AspGlnAlaAspTyrGlnAlaGlyArgArgArgLeuLeuMetGlnGlnGlnGlyGly	319
QY	1081	CGGGCGCCGACAGAGGCGCTCC	1101
Db	320	ArgArgProThrGlnAlaSer	326
RESULT 2			
DED2_MOUSE		STANDARD;	PRT; 330 AA.
AC	08QZV0;	08JZV1;	
DT	05-JUN-2004	(Rel. 44, Created)	
DT	05-JUL-2004	(Rel. 44, Last sequence update)	
DT	05-JUL-2004	(Rel. 44, Last annotation update)	
DE	DNA-binding deach effector domain-containing protein 2 (DED-containing protein FLAME-3).		
GN	Name=Ded2; Synonyms=Flame3;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RC	STRAIN=C57BL/6J		
RX	MEDLINE=21961615;	PubMed=11965497;	DOI=10.1038/sj/cdd/4401038;
RA	Zhan Y., Hegge R., Srinivasula S.M., Fernandes-Alnemri T., Alnemri E.S.;		
RT	"Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the TRIF102 subunit of human transcription factor TRIC".		
RL	Cell Death Differ. 9:439-447(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RX	PubMed=12527898;	DOI=10.1038/sj.onc.1206099;	
RA	Alcivar A., Hu S., Tang J., Yang X.;		
RL	"DEDD and DEDD2 associate with caspase-8/10 and signal cell death.";		
RN	Oncogene 22:291-297(2003).		
RN	[3]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	TISSUE=Retina;		
RX	MEDLINE=22388257;	PubMed=12477937;	DOI=10.1073/pnas.242603899;
RA	Straubeberg R.L., Feingold E.A., Grouse L.H., Derye J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Hopkins S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Dickens R.F., Jordan H., Moore T., Max S.I., Wang J., Heif F., Ditschenko L., Marusina K., Farmer A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein W.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,		

RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600.
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Ishi Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kaibiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai U.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kita A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cortex;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiroka T., Hirozane T.,
RA Hori F., Imotani K., Ishi Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK043308; BAC31639.1; -.
DR MGD; MGI:1914629; Dded2.
DR GO; GO:0042981; P:regulation of apoptosis; TAS.
KW DNA-binding
SQ SEQUENCE 168 AA; 17242 MW; F75D71DC489D2425 CRC64;

Alignment Scores:
Pred. No.: 1,45e-12 Length: 168
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.71% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x Q8BRM9 (1-168)
QY 706 ACCTGTGACATCCGGCTCCGGGTTTCAGACAGACTGCGAGCATGGCCAGCCTTGAG 765
Db 83 ThCysAspIleArgLeuArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGlu 102
QY 766 CAGGGCGTG 774
Db 103 GlnGlyVal 105

RESULT 4
Q919M3 PRELIMINARY; PRT; 369 AA.
ID 0919M3;
AC 0919M3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
Ded1.
GN Name=ded1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;

RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish."
RL Cell Death Differ. 7:509-510(2000).
RN [5]
RP EMBL; AF222226; AAF66963.1; -.
DR ZFIN; ZDB-GENE-000616-2; ded1.
DR GO; GO:0005515; F:protein binding; IEA.
DR GO; GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro; IPR011029; DEATH_1like.
DR InterPro; IPR001875; DED.
DR Pfam; PF01335; DED_1.
DR PROSITE; PS50168; DED; 1.
SQ SEQUENCE 369 AA; 42244 MW; 73B09E9E17ECC247 CRC64;

Alignment Scores:
Pred. No.: 0.000754 Length: 369
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.26% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x Q919M3 (1-369)
QY 700 AAAGTGAACCTGTGACATCCGGCTCCGGGTTTCAGACAGATAC 741
Db 235 LysValThrCysAspIleArgLeuArgValArgAlaGluTyr 248

RESULT 5
DED_HUMAN
ID DEDD_HUMAN STANDARD; PRT; 318 AA.
AC 075618; O60737;
DT 28-FEB-2003 (Ref. 41, Created)
DT 28-FEB-2003 (Ref. 41, Last sequence update)
DT 05-JUL-2004 (Ref. 44, Last annotation update)
DE Death effector domain-containing protein (Death effector domain-
DE containing testicular molecule) (DEDProl) (F1DED-1) (KE05).
GN Name=DEDD; Synonyms=DEDProl, DEFT;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98447539; PubMed=9774341; DOI=10.1093/emboj/17.20.5974;
RA Stegh A.H., Schickling O., Ehret A., Scalfidi C., Peterhansel C.,
RA Hofmann T.G., Grunmt I., Krammer P.H., Peter M.E.;
RT "DEDD, a novel death effector domain-containing protein, targeted to
RT the nucleolus."
RL EMBO J. 17:5974-5986(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISSUE=Testis;
RX MEDLINE=99049260; PubMed=9832420; DOI=10.1210/en.139.12.4839;
RA Leo C.P., Heu S.Y., McGee E.A., Salanova M., Haneh A.J.W.;
RT "DEFT, a novel death effector domain-containing molecule predominantly
RT expressed in testicular germ cells."
RL Endocrinology 139:4839-4848(1998).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
AC Pan G.;
DT "F1DED-1, a novel molecule with a DED-like domain."
DT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Thome M., Tschopp J.;
RT "DEDProl, a novel DED-containing protein."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Dendritic cell;
RA Zhao Z., Huang X., Li N., Zhu X., Cao X.;
RT "A novel gene from human dendritic cell."

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Bone marrow, and Placenta;
 RX MEDLINE=22288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Steinmetz C.M., Shat N.K.,
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.U., Ustin T.B., Toshiyuki S., Carninci P., Pearce C.T.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP FUNCTION, AND INTERACTIONS WITH KRT8; KRT18 AND CASP3.
 RX MEDLINE=2220288; PubMed=12235123; DOI=10.1083/jcb.200112124;
 RA Lee J.C., Schickling O., Stegh A.H., Oshima R.G., Dinsdale D.,
 RA Cohen G.M., Peter M.E.,
 RT "DED regulates degradation of intermediate filaments during
 RT apoptosis.";
 RL J. Cell Biol. 158:1051-1066(2002).
 RN [8]
 RP INTERACTION WITH GFF3.
 RX MEDLINE=21961615; PubMed=11965497; DOI=10.1038/sj/cdd.4401038;
 RA Zhan Y., Hegde R., Srinivasula S.M., Fernandes-Alnemri T.,
 RA Alnemri E.S.;
 RT "Death effector domain-containing proteins DEDD and FLAME-3 form
 RT nuclear complexes with the TPIC102 subunit of human transcription
 RT factor IIIC.";
 RL Cell Death Differ. 9:439-447(2002).
 RN [9]
 RP INTERACTIONS WITH CASP8 AND CASP10.
 RX PubMed=12527898; DOI=10.1038/sj.onc.1206099;
 RA Alcivar A., Hu S., Tang J., Yang X.;
 RT "DED and DEDD2 associate with caspase-8/10 and signal cell death.";
 RL Oncogene 22:291-297(2003).
 CC -1- FUNCTION: A scaffold protein that directs CASP3 to certain
 CC substrates and facilitates their ordered degradation during
 CC apoptosis. May also play a role in mediating CASP3 cleavage of
 CC KRT18. Regulates degradation of intermediate filaments during
 CC apoptosis. May play a role in the general transcription machinery
 CC in the nucleus and might be an important regulator of the activity
 CC of GFF3. Inhibits DNA transcription in vitro (By similarity).
 CC -1- SUBUNIT: Interacts with CASP8, CASP10, KRT8, KRT18, CASP3 and
 CC FADD. Homodimerizes and heterodimerizes with DEDD2.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
 CC during CD95-mediated apoptosis where it is localized in the
 CC nucleolus (by similarity). Following apoptosis induction, the mono
 CC and/or dimerization form increases and forms filamentous
 CC structures that colocalize with KRT8 and KRT18 intermediate
 CC filament network in simple epithelial cells.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=O75618-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=O75618-2; Sequence=VSP_003846;
 CC Note=No experimental confirmation available;
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
 CC testis.
 CC -1- PFM: Exists predominantly in a mono- or dimerized form.

CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; AF083236; AAC33105.1; -
 DR EMBL; AF100341; AAD16414.1; -
 DR EMBL; AF043733; AAC80280.1; -
 DR EMBL; AF010973; CA09445.1; -
 DR EMBL; AF064605; AAC17110.3; -
 DR EMBL; BC016724; AAC16724.1; -
 DR EMBL; BC013910; AAH13910.1; -
 DR Genbank; HGNC:2755; DEDD.
 DR H-InvDB; HIX0001231; -
 DR MIM; 606841; -
 DR GO; GO:0005737; Cytoplasm; ISS.
 DR GO; GO:0005730; Cnucleolus; ISS.
 DR GO; GO:0003677; P:DNA binding; ISS.
 DR GO; GO:0006917; P:induction of apoptosis; ISS.
 DR GO; GO:0008625; P:induction of apoptosis via death domain rec.; TAS.
 DR GO; GO:0016481; P:negative regulation of transcription; ISS.
 DR GO; GO:0007283; P:spermatogenesis; TAS.
 DR InterPro; IPR011029; DEATH_1like.
 DR InterPro; IPR01875; DED.
 DR Pfam; PF01335; DED; 1.
 DR SMART; SM00031; DED; 1.
 DR PROSITE; PS0168; DED; 1.
 DR Alternative splicing; Apoptosis; DNA-binding; Nuclear protein;
 KW Repressor; Transcription regulation.
 FT DOMAIN 25 103 DED.
 FT VARSPPLIC 194 194 DED.
 FT D -> GEEIQGFQMSRLGEYKELLGHVAVYAIQY (in
 FT isoform 2).
 FT /FTId=VSP_003846.
 FT P -> L (in Ref. 5).
 FT CONFLICT 13 13
 FT SEQUENCE 318 AA; 36794 MW; FF9D5FP9B61F6BB6 CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 0.00723 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: 1 Gaps: 0
 US-10-030-271-3 (1-1893) x DEDD HUMAN (1-318)
 QY 706 ACCGTGACATCCGGCTCCGGGTTCCAGCAGACTACTGC 744
 Db 192 ThrcyaspilleargleuargvalargalagluTYrCys 204
 RESULT 6
 DEDD MOUSE STANDARD; PRT; 318 AA.
 ID DEDD MOUSE
 AC 0921L3; O7TOH8; Q9R227;
 DT 28-FEB-2003 (Rel. 41; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 05-JUL-2004 (Rel. 44; Last annotation update)
 DE Death effector domain-containing protein (DEDDpro1).
 GN Name=DedD;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RX MEDLINE=98447599; PubMed=9774341; DOI=10.1093/emboj/17.20.5974;
 RA Stegh A.H., Schickling O., Ehret A., Scaffidi C., Peterhaensel C.,

RA Hofmann T.G., Grunmt I., Krammer P.H., Peter M.E.;
 RT "DEDD, a novel death effector domain containing protein, targeted to
 the nucleolus.";
 RL EMO J. 17:5974-5986 (1998).
 RN [2]
 RA SEQUENCE FROM N.A.
 RA Thome M., Tschopp J.;
 RT "DEDPOL, a novel DED-containing protein.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N-3; TISSUE=Colon, and Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.M.,
 RA Krausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
 RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitoki S., Cantin P., Prange C.,
 RA Raba S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay D.J., Holys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeney R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Butterch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: A scaffold protein that directs CASP3 to certain
 CC substrates and facilitates their ordered degradation during
 CC apoptosis. May also play a role in mediating CASP3 cleavage of
 CC KRT18. Regulates degradation of intermediate filaments during
 CC apoptosis. May play a role in the general transcription machinery
 CC in the nucleus and might be an important regulator of the activity
 CC of GTF3C3 (By similarity). Inhibits DNA transcription in vitro.
 CC -1- SUBUNIT: Interacts with CASP8, CASP10, KRT8, KRT18, CASP3 and
 CC FADD. Homodimerizes and heterodimerizes with DEDD2 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
 CC during CD95-mediated apoptosis where it is localized in the
 CC nucleoli. Following apoptosis induction, the mono and/or
 CC dimerization form increases and forms filamentous structures
 CC that colocalize with KRT8 and KRT18 intermediate filament network
 CC in simple epithelial cells (By similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- PTM: Exists predominantly in a mono- or dimerized form.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>;
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AJ011386; CA09604.1; -;
 CC EMBL: AF100342; AAD16415.1; -;
 CC EMBL: BC023668; AAH23668.1; -;
 CC EMBL: BC054445; AAH54445.1; -;
 CC EMBL: MGI:1333874; DEDD.
 CC GO: GO:0005737; C:cytoplasm; IDA.
 CC GO: GO:0005730; C:nucleolus; IDA.
 CC GO: GO:0006377; F:DNA binding; IDA.
 CC GO: GO:0006917; P:induction of apoptosis; IDA.
 CC GO: GO:0016481; P:negative regulation of transcription; IDA.
 CC Interpro: IPR011029; DEATH like.
 CC Interpro: IPR001875; DED.

DR Pfam; PF01335; DED; 1.
 DR SMART; SM00031; DED; 1.
 DR PROSITE; PS50168; DED; 1.
 KM Apoptosis; DNA-binding; Nuclear protein; Repressor;
 KM Transcription regulation.
 FT DOMAIN 25 103 DED.
 FT CONFLICT 237 237 K -> N (in Ref. 2).
 FT CONFLICT 316 316 A -> V (in Ref. 3; AAH54445).
 SO SEQUENCE 318 AA; 36805 MW; C9A31DFC4C0B57CA CRC64;
 Alignment Scores:
 Pred. No.: 0.00723 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: Gaps: 0
 US-10-030-271-3 (1-1883) x DEDD_MOUSE (1-318)
 QY 706 ACCGTGACATCCGGCTCCGGCTTCGACGACAGTACTGC 744
 Db 192 ThrcysaplleargueuargvalArgAlaGluTyrCys 204
 RESULT 7
 ID DEDD RAT STANDARD; PRT; 318 AA.
 AC Q922K0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Death effector domain-containing protein (Death effector domain-
 DE containing testicular molecule).
 GN Name=Dead; Synonyms=Defc;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Testis;
 RX MEDLINE=99049260; PubMed=9832420; DOI=10.1210/en.139.12.4839;
 RA Leo C.P., Hsu S.Y., McGee E.A., Salanova M., Hsueh A.J.W.;
 RT "DEFT, a novel death effector domain-containing molecule predominantly
 RT expressed in testicular germ cells.";
 RL Endocrinology 139:4839-4848(1998).
 CC -1- FUNCTION: A scaffold protein that directs CASP3 to certain
 CC substrates and facilitates their ordered degradation during
 CC apoptosis. May also play a role in mediating CASP3 cleavage of
 CC KRT18. Regulates degradation of intermediate filaments during
 CC apoptosis. May play a role in the general transcription machinery
 CC in the nucleus and might be an important regulator of the activity
 CC of GTF3C3. Inhibits DNA transcription in vitro (By similarity).
 CC -1- SUBUNIT: Interacts with CASP8, CASP10, KRT8, KRT18, CASP3 and
 CC FADD. Homodimerizes and heterodimerizes with DEDD2 (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Translocated to the nucleus
 CC during CD95-mediated apoptosis where it is localized in the
 CC nucleoli. Following apoptosis induction, the mono and/or
 CC dimerization form increases and forms filamentous structures
 CC that colocalize with KRT8 and KRT18 intermediate filament network
 CC in simple epithelial cells (By similarity).
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in
 CC testis. Within the testis, highly expressed in germ cells but not
 CC expressed in Sertoli cells.
 CC -1- DEVELOPMENTAL STAGE: First detected in 20-day-old animals. Reaches
 CC a peak at 30 days.
 CC -1- PTM: Exists predominantly in a mono- or dimerized form.
 CC -1- SIMILARITY: Contains 1 death effector (DED) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).

DR EMBL: AF053162; AAC08287.1; -
DR GO: GO:0005737; C:cytoplasm; ISS.
DR GO: GO:0005730; C:nucleolus; ISS.
DR GO: GO:0003677; F:DNA binding; ISS.
DR GO: GO:0006917; P:induction of apoptosis; ISS.
DR GO: GO:0016481; P:negative regulation of transcription; ISS.
DR InterPro: IPR011029; DEATH_like.
DR InterPro: IPR001875; DED.
DR Pfam: PF01335; DED; 1.
DR SMART: SM00031; DED; 1.
DR PROSITE: PS50168; DED; 1.
KM Apoptosis: DNA-binding; Nuclear protein; Repressor;
PT Transcription regulation.
FT DOMAIN 25 103 DED.
SQ SEQUENCE 318 AA; 36847 MW; B8751791F66A03DE CRC64;

Alignment Scores:
Pred. No.: 0.00723 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x DED_RAT (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGTTTCGAGCAGACTGTC 744
Db 192 Thrcyasp1learg1euarValargAlaglutyrcys 204

RESULT 8

Q6DHN2 PRELIMINARY; PRT; 404 AA.
ID Q6DHN2
AC Q6DHN2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:92202.
GN Name=zgc:92202;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxID=7955;
RN NCB1
RP
RC
RS
RT
RT and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RA Strausberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC075935; AAH75935.1; -
DR GO: GO:0005515; F:protein binding; IEA.
DR GO: GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro: IPR011029; DEATH_like.
DR InterPro: IPR001875; DED.
DR Pfam: PF01335; DED; 1.
DR SMART: SM00031; DED; 1.
DR PROSITE: PS50168; DED; 1.
SQ SEQUENCE 404 AA; 44888 MW; 3C949DB8B3B078B1A CRC64;

Alignment Scores:
Pred. No.: 0.00701 Length: 404
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x Q6DHN2 (1-404)

QY 706 ACCTGTGACATCCGGCTCCGGTTTCGAGCAGACTGTC 744
Db 267 Thrcyasp1learg1euarValargAlaglutyrcys 279

RESULT 9

Q6GNZ8 PRELIMINARY; PRT; 243 AA.
ID Q6GNZ8
AC Q6GNZ8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MGC80767 protein.
GN Name=MGC80767;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoes; Pipidae;
OC Xenopodidae; Xenopus.
NCBI_TaxID=6355;
RN NCB1
RP
RC
RS
RT
RT and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

```

RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT Initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC073351; AAF73351.1; -
DR GO: GO:0005515; F:protein binding; IEA.
DR GO: GO:0042881; P:regulation of apoptosis; IEA.
DR InterPro: IPR011029; DEATH_like.
DR Pfam: PF01335; DED; 1.
DR PROSITE: PS50168; DED; 1.
SQ SEQUENCE 243 AA; 27767 MW; 57A2B89CFBD3E0BC CRC64;

Alignment Scores:
Pred. No.: 0.661 Length: 243
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.77% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x Q6GNZ8 (1-243)

QY 181 TACTACGGAGTGTGCTGCTTACCCGATGTC 213
Db 21 TyrtYrGlYmEtSerLeuH1sArgMcPhe 31

RESULT 10
Q91W13 PRELIMINARY; PRT; 218 AA.
ID Q91W13
AC Q91W13;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Interleukin 18 binding protein.
GN Name=0541;
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OC NCBI_TaxID=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20231932; PubMed=10769064;
RT Smith V.P., Bryant N.A., Alcamí A.;
RT "Ectromelia, vaccinia and cowpox viruses encode secreted interleukin
RT 18 binding proteins."
RL J. Gen. Virol. 81:1223-1230(2000).
DR EMBL: AJ271163; CAB89814.1; -
SQ SEQUENCE 218 AA; 23853 MW; 3FPA9C24DD24E75 CRC64;

Alignment Scores:
Pred. No.: 6.31 Length: 218
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x Q91W13 (1-218)

QY 622 GGTGCCAGACGGCGCGGAGAGGGGCCCA 651
Db 160 G1yAlaArgArgArgArgArgG1yAlaPro 169

RESULT 11
Q98222 PRELIMINARY; PRT; 235 AA.
AC Q98222;
DT 01-FEB-1997 (TREMBlrel. 02, Created)

RT "Genome sequence of a human tumorigenic poxvirus: prediction of
RT specific host response-evasion genes."
RL Science 273:813-816(1996).
DR EMBL: U60315; AAC55182.1; -
DR PIR: T30656; T30656.
SQ SEQUENCE 235 AA; 25209 MW; BA47745C68608889 CRC64;

Alignment Scores:
Pred. No.: 6.25 Length: 235
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x Q98222 (1-235)

QY 622 GGTGCCAGACGGCGCGGAGAGGGGCCCA 651
Db 156 G1yAlaArgArgArgArgArgG1yAlaPro 165

RESULT 12
Q6DHV2 PRELIMINARY; PRT; 244 AA.
ID Q6DHV2
AC Q6DHV2;
DT 25-OCT-2004 (T-EMBlrel. 28, Created)
DT 25-OCT-2004 (T-EMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (T-EMBlrel. 28, Last annotation update)
DE Dedd protein.
GN Name=dedd1;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole;
RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Matsushita K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ushed T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Miall S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalms D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Whole:
RA Straussberg R.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC075865; AA075865.1; -.
DR GO: GO:0005515; P:protein binding; IEA.
DR GO: GO:0042981; P:regulation of apoptosis; IEA.
DR InterPro: IPR011029; DEATH_like.
DR InterPro: IPR011875; DED.
DR Pfam: PF01335; DED; 1.
DR PROSITE: PSS0168; DED; 1.
SQ SEQUENCE 244 AA; 28051 MW; 043225AEAD79527 CRC64;

Alignment Scores:
Pred. NO.: 6.22 Length: 244
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservatide: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q6DHV2 (1-244)
Oy 322 GAGCTCCTGCTGAGCTGAGCGCCGCGG 351
|||
Db 113 GtluenuleuGtluenuGtluarGagGly 122

RESULT 13
O9F1E2
ID O9F1E2 PRELIMINARY; PRT; 581 AA.
AC O9F1E2; O9ZH33;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcriptional regulator NifA.
GN Name:nifA;
OS Acetobacter diazotrophicus (Gluconacetobacter diazotrophicus).
OC Bacteriia; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
CC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=33996;
RN RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAL5;
RA Sevilla M.Q., de Oliveira A., Baldani I., Kennedy C.;
RT "Contributions of the bacterial endophyte Acetobacter diazotrophicus
RT to sugarcane nutrition: A preliminary study.";
RL Symbiosis 25:181-192(1998).
RN RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PAL5;
RA Teixeira K.R.S., Whelling M., Morgan T., Gallier R., Zellerman E.,
RA Baldani I.I., Kennedy C., Meletzus D.;
RT "Molecular analysis of the chromosomal region encoding the nifA and
RT nifB genes of Acetobacter diazotrophicus.";
RL FEMS Microbiol. Lett. 176:301-309(1999).
RN RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=PAL5;
RX MEDLINE=20545459; Pubmed=11092875;
RX DOI=10.1128/JB.182.24.7088-7091.2000;
RA Lee S., Reth A., Meletzus D., Sevilla M., Kennedy C.;
RT "Characterization of a major cluster of nif, fix, and associated genes
RT in a sugarcane endophyte, Acetobacter diazotrophicus.";
RL J. Bacteriol. 182:7088-7091(2000).
RN RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=PAL5;
RA Meletzus D., Reth A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=PAL5;
RA Meletzus D., Reth A.;

```

RL Submitted (MAR-2001) to the EMBL/Genbank/DBD databases.
 CC -1- SIMILARITY: Contains 1 sigma=54 factor interaction ATP-binding
 domain.
 DR EMBL, AF030414; AAC9364.2; -.
 DR GO, GO:0005524; F:ATP binding; IEA.
 DR GO, GO:0000166; F:nucleotide binding; IEA.
 DR GO, GO:0003700; F:transcription factor activity; IEA.
 DR GO, GO:0016563; F:transcriptional activator activity; IEA.
 DR GO, GO:0009399; F:nitrogen fixation; IEA.
 DR GO, GO:0000160; P:two-component signal transduction system (p. . .; IEA
 DR InterPro, IPRO03593; AAA ATPase.
 DR InterPro, IPRO08931; FIS-like.
 DR InterPro, IPRO03018; GAF.
 DR InterPro, IPRO02197; HTH_Fis.
 DR InterPro, IPRO10113; NIFA.
 DR InterPro, IPRO02078; Sig54_interact.
 DR Pfam, PF01590; GAF. 1.
 DR Pfam, PF02954; HTH_8. 1.
 DR Pfam, PF00158; Sigma54_activat; 1.
 DR PRINTS, PRO1590; HTHFIS.
 DR SMART, SM00382; AAA; 1.
 DR SMART, SM00065; GAF; 1.
 DR TIGRFAMs, TIGR01199; HTH_fis; 1.
 DR TIGRFAMs, TIGR01817; nifa; 1.
 DR PROSITE, PS00675; SIGMA54_INTERACT_1; 1.
 DR PROSITE, PS00676; SIGMA54_INTERACT_2; 1.
 DR PROSITE, PS00688; SIGMA54_INTERACT_3; 1.
 DR PROSITE, PSS0045; SIGMA54_INTERACT_4; 1.
 DR ATP-binding; DNA-binding; Transcription; Transcription regulation.
 QO SEQUENCE 581 AA; 62322 MW; 17445DC3235FFC91 CRC64;

Alignment Scores:	5.56	Length:	581
Pred. No.:	10.00	Matches:	10
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	1.61%	Indels:	0
Query Match:	2	Gaps:	0
DB:			
US-10-030-271-3 (1-1883) x Q9F1E2 (1-581)			
QY	1044 CCGGCGCGCGCTTGTGATGAGGAGGA	1073	
DB	203 ProAlaProProValAlaAspGlyGly	212	
RESULT 14			
Q91TW1			
ID	Q91TW1	PRELIMINARY;	PRT; 729 AA.
AC	Q91TW1;		
DT	01-DEC-2001 (TrEMBL:rel. 19, Created)		
DT	01-DEC-2001 (TrEMBL:rel. 19, Last sequence update)		
DT	01-OCT-2003 (TrEMBL:rel. 25, Last annotation update)		
DE	T2.		
OS	Tupaiaid herpesvirus 1 (strain 1) (TnHV-1) (Herpesvirus tupaia (strain 1)).		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Betaherpesvirinae.		
OC	NCBI_TaxID=10397;		
CX	(1)		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=2;		
RC			
RX	MEDLINE=21211637; PubMed=11312357;		
RX	DOI=10.1128/JVI.75.10.4854-4870.2001;		
RA	Bahr U., Darai G.,		
RT	"Analysis and characterization of the complete genome of tupaia (tree shrew) herpesvirus.";		
RT	J. Virol. 75:4854-4870(2001).		
RL	EMBL; AF281817; AAKS7026.1; -.		
DR	PFam; PF02393; US22; 1.		
SQ	SEQUENCE 729 AA; 81191 MW; D5FDE57AAB3EAB5 CRC64;		
Alignment Scores:	5.4	Length:	729
Pred. No.:			

Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q91TW1 (1-729)

QY 347 CGGCGCTCCAGCTCCAGCAGAGCTTAG 318
 DB 465 ArgArgSerSerSerArgSerArg 474

RESULT 15

IE18 PRIVIF STANDARD; PRT; 1461 AA.

AC P1675;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Immediate-early protein IE180.

GN Name=IE;
 OS Pseudorabies virus (strain Indiana-Punkhauser / Becker) (PRV).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.

NCBI_TaxID=31523;

RP SEQUENCE FROM N.A.
 RX MEDLINE=89315207; PubMed=2546124;

RA Cheung A.K.;
 RT "DNA nucleotide sequence analysis of the immediate-early gene of
 RT Pseudorabies virus.";

RL Nucleic Acids Res. 17:4637-4646(1989).
 RN [2]

REVISIONS.

RA Cheung A.K.;
 RL Submitted (NOV-1989) to the EMBL/GenBank/DBJ databases.

CC - FUNCTION: This IE protein is a multifunctional protein capable of
 CC migrating to the nucleus, binding to DNA, trans-activating other
 CC viral genes, and autoregulating its own synthesis.

CC - SUBCELLULAR LOCATION: Nucleus of infected cells.
 CC - PTM: A long stretch of serine residues may be a major site of
 CC phosphorylation.

CC - SIMILARITY: Belongs to the herpesviruses ICP4/IE140/IE180 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL, X15120; CAJ3214.1; -.

DR PIR; S04713; EDBEIR.

DR InterPro; IPR005205; Herpes ICP4_C.

DR InterPro; IPR005206; Herpes ICP4_N.

DR Pfam; PF03585; Herpes_ICP4_C_1.

DR Pfam; PF03584; Herpes_ICP4_N_1.

KM DNA-binding; Early protein; Nuclear protein; Phosphorylation;
 KM Trans-acting factor; Transcription regulation.

PT DOMAIN 390 405 Poly-Ser.

FT DOMAIN 958 966 Poly-Ser.

SO SEQUENCE 1461 AA; 149833 MW; 7F31E7ABE403B208 CRC64;

Alignment Scores:

Pred. No.: 4.93 Length: 1461
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.62% Indels: 0
 DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x IE18_PRIVIF (1-1461)

QY 659 GGTGGCGCTGGGCCCCCTCTCCGCCCG 630
 DB 1047 GYAlaIaIaGlyAlaProIeuAlaArgAArg 1056

RESULT 16

ID Q65553 PRELIMINARY; PRT; 3247 AA.

AC Q65553;
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)

DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)

DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)

DE UL36.

OS Bovine herpesvirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.

NCBI_TaxID=10320;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Cooper;

RA Schwyzer M., Vlack C., Lowery D.E., Bello L.J., Meyer G., Misra V.,
 RA Thiry E., Paces V.;

RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z78205; CAB01605.1; -.

DR InterPro; IPR006928; Herpes_UL36_N.

DR InterPro; IPR005210; Herpes_UL36.

DR Pfam; PF04843; Herpes_UL36_N_1.

DR Pfam; PF03586; Herpes_UL36; 1.

DR SEQUENCE 3247 AA; 332190 MW; 3AEAA72F8F001F6A CRC64;

US-10-030-271-3 (1-1883) x Q65553 (1-3247)

QY 87 CCGGCTTCGAGCTGTTCGCTCCCTCC 116

DB 2786 ProGlySerGluIeuValProProProSer 2795

RESULT 17

ID Q77CD4 PRELIMINARY; PRT; 3247 AA.

AC Q77CD4;
 DT 05-JUL-2004 (T-EMBLrel. 27, Created)

DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)

DE Very large virion protein (tegument).

GN Name=UL36;

OS Bovine herpesvirus type 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.

NCBI_TaxID=79889;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97164286; PubMed=9010999; DOI=10.1016/S0378-1135(96)01235-7;
 RA Schwyzer M., Stuyger D., Vogt B., Lowery D.E., Simard C.,
 RA Laboisliere S., Misra V., Vlack C., Paces V.;

RT "Gene contents in a 31-kb segment at the left genome end of bovine
 RT herpesvirus-1.";

RT Vet. Microbiol. 53:67-77(1996).

```

RN [2]
RP SEQUENCE FROM N.A.
RA Schwyzner M., Vlcek C., Lowery D.E., Bello L.J., Meyer G., Misra V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Schwyzner M., Paces V., Letchworth G.J., Misra V., Bunk H.J.,
RA Lowery D.E., Simard C., Bello L.J., Thilly E., Vlcek C.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; A004801; CAA06097.1; -
DR InterPro; IPR006928; Herpes teg. N.
DR InterPro; IPR005210; Herpes teg. N.
DR Pfam; PF04843; Herpes_teg_N; 1.
DR Pfam; PF03586; Herpes_UL36; 1.
SQ SEQUENCE 3247 AA; 332190 MW; 3AEAA72F8F001F6A CRC64;

Alignment Scores:
Pred. No.: 4.45 Length: 3247
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x Q77CD4 (1-3247)

QY 87 CCGGGTTCGAGCTGTTCGCCCTCCCTCC 116
Db 2786 Progl|ySerg|lueu|val|Pro|Pro|Proser 2795

RESULT 18
Q8W615 PRELIMINARY; PRT; 110 AA.
AC Q8W615;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein orf42.
DE Name=orf42;
OS Sinchizobium mellici phage PBC5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales.
OX NCBI_TaxID=179237;
RN [1]
RP SEQUENCE FROM N.A.
RA Schumelster S.A., Krol J.E., Vorhoeelter F.-J., Skorupska A.M.,
RA Lotz W.;
RT "Sequence of the genome of Sinchizobium mellici bacteriophage
RT PBC5.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Schumelster S.A., Krol J.E., Vorhoeelter F.-J., Skorupska A.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF449724; AA149638.1; -
DR Hypothetical protein.
SQ SEQUENCE 110 AA; 12194 MW; 1EF7398F8CF302E8 CRC64;

Alignment Scores:
Pred. No.: 64.8 Length: 110
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x Q8W615 (1-110)

QY 338 AGCTCCAGCAGAGCTCTAGCCGCTG 312
Db 100 SerSerSerArgSerSerArgProleu 108

RESULT 19
Q8CS18
```

```

ID Q8CS18 PRELIMINARY; PRT; 115 AA.
AC Q8CS18;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Mus musculus adult male olfactory brain cDNA, RIKEN full-length
DE enriched library, clone:6430539E12 product:hypothetical protein, full
DE insert sequence.
GN Name=2610204G22Rlk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carrinci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carrinci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carrinci P.,
RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Olfactory brain;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carrinci P.,
RA Fukuda S., Furuta M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai C., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
```

DR EMBL; AK078249; BAC37192.1; -
 DR MGD; MGI:1917698; 2610204G22R1K.
 KM Hypothetical protein.
 SQ SEQUENCE 115 AA; 12569 MW; BBEBC246D9D4C52 CRC64;

Alignment Scores:

Pred. No.: 64.4 Length: 115
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q8C5I8 (1-115)

Qy 237 CGAGTCGAGCTGAGCTCTGAGCCT 263
 Db 95 ArgvalaArgAlaGlyAlaProGlyLeu 103

RESULT 20

Q9Y9Y9 PRELIMINARY; PRT; 127 AA.
 AC Q9Y9Y9;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein APE2150.
 GN OrderedLocustNames=APE2150;
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
 OC Desulfurococcaceae; Aeropyrum.
 NCBI_TaxID=56636;
 RN NCBI_TaxID=56636;
 RP SEQUENCE FROM N.A.

RC STRAIN=K1;
 RX MEDLINE=9310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamizaki M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Nomura N., Sako Y., Kikuchi H.,
 RA Nakamura Y., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1."
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000063; BAB81161.1; -
 DR PIR; A72522; A72522.
 KM Complete proteome; Hypothetical protein.
 SQ SEQUENCE 127 AA; 13517 MW; 1C28D0E67B34434C CRC64;

Alignment Scores:

Pred. No.: 63.6 Length: 127
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q9Y9Y9 (1-127)

Qy 469 AGCTATGCACTTCACGCTCTTCAAG 495
 Db 95 SerTyrGlyThrSerSerSerTyrLeu 103

RESULT 21

Q67VU8 PRELIMINARY; PRT; 148 AA.
 ID Q67VU8;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein OSUNB0023122.20 (Hypothetical protein
 DE P0530H05.35).

GN Name=OSUNB0023122.20; Synonyms=P0530H05.35;
 OS Oryza sativa (Japonica cultivar-group).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriartoideae; Oryzaceae; Oryza.
 NCBI_TaxID=39947;
 RN NCBI_TaxID=39947;
 RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
 RT clone:OSUNB0023122.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
 RT clone:P0530H05.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004733; BAD37721.1; -
 DR EMBL; AP003541; BAD37396.1; -
 KM Hypothetical protein.
 SQ SEQUENCE 148 AA; 14721 MW; 7CDBC2564E30DEF9 CRC64;

Alignment Scores:

Pred. No.: 62.3 Length: 148
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q67VU8 (1-148)

Qy 811 CCAGTCGCGCGCCGAGCCTGAGCCT 785
 Db 29 ProAlaAlaAlaProAlaProGlyAla 37

RESULT 22

Q6ZLD4 PRELIMINARY; PRT; 201 AA.
 AC Q6ZLD4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Cyclophilin-like protein.
 GN Name=OJ1014_E09.6; Synonyms=OJ1361_E02.124;
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriartoideae; Oryzaceae; Oryza.
 NCBI_TaxID=39947;
 RN NCBI_TaxID=39947;
 RP SEQUENCE FROM N.A.

RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OJ1014_E09.6";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
 RT clone:OJ1361_E02.1";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP003800; BAC83031.1; -
 DR EMBL; AP003826; BAD30334.1; -
 DR InterPro; IPR00194; ATPase_a/centr.
 DR PROSITE; PS00152; ATPase ALPHA BETA; UNKOWN 1.
 SQ SEQUENCE 201 AA; 23173 MW; D31CD1567660777D CRC64;

Alignment Scores:

Pred. No.: 59.9 Length: 201
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0
US-10-030-271-3 (1-1883) x Q6ZLD4 (1-201)

QY 347 CGCGCTCCGCTCCAGCAGGAGCTCT 321
DB 182 ArgArgSerSerSerArgSerSer 190

RESULT 23
Q67IV3 PRELIMINARY; PRT; 222 AA.
AC Q67IV3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein OSUNBa0048L03.36.
GN Name=OSUNBa0048L03.36;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Fujisawa M.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC
clone:OSUNBa0048L03.36";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP007205; BAD38588.1; -.
KW Hypothetical protein.
SQ SEQUENCE 222 AA; 24596 MW; 95CSB33F65534936 CRC64;

Alignment Scores:
Pred. No.: 59.2 Length: 222
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x Q67IV3 (1-222)

QY 619 GGTGTCGACGCGCGGAGAGG 645
DB 85 G1YGLYAlaArgArgArgArgGly 93

RESULT 24
Q7Q6G3 PRELIMINARY; PRT; 227 AA.
AC Q7Q6G3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP6626 (Fragment).
GN Name=agCG48835; ORFNames=ENSA00000015017;
OS Anopheles gambiae str. PE8T.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PE8T;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
EMBL; AAB01008960; EAA11634.1; -.
DR GO; GO:0005783; C:cytoplasmic reticulum; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR008958; Transglut_C.

DR InterPro; IPR008856; TRAP_beta.
DR Pfam; PF05753; TRAP_beta; 1.
FT NON_TER 1
SQ SEQUENCE 227 AA; 24921 MW; 2ADF4D013495BD59 CRC64;

Alignment Scores:
Pred. No.: 59 Length: 227
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x Q7Q6G3 (1-227)

QY 211 TTCGAGTGTGGCGGCGCACTGACC 237
DB 109 PheGluValValGlyGlyGlnLeuThr 117

RESULT 25
Q6MBX0 PRELIMINARY; PRT; 241 AA.
AC Q6MBX0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderediocName=pcl205;
OS Parachlamydia sp. (strain UWE25) (subsp. Acanthamoeba sp.).
OC Bacteria; Chlamydiae; Chlamydiales; Parachlamydiaceae; Parachlamydia.
OX NCBI_TaxID=264201;
RN [1]
RP SEQUENCE FROM N.A.
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhoid U.,
RA Rattmann B., Brandt P., Nyakatura G.J., Drooge M., Frishman D.,
RA Rattei T., Mewes H.-W., Wagner M.;
RT "genome sequence of an amoeba symbiont and its use for reconstructing
the evolutionary history of chlamydiae";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX908798; CAF23929.1; -.
DR InterPro; IPR009058; Wing_hix_DNA_bnd.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 241 AA; 27948 MW; DD3BCAF404961E3 CRC64;

Alignment Scores:
Pred. No.: 58.5 Length: 241
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x Q6MBX0 (1-241)

QY 316 GGCCTAGACTCTCTGCTGAGCTGGAG 342
DB 166 GlyLeuGlnLeuLeuGlnLeuGln 174

RESULT 26
Q9FTL5 PRELIMINARY; PRT; 243 AA.
AC Q9FTL5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE P0034C11.5 protein (OSUNBa0004G10.24 protein).
GN Name=P0034C11.5; Synonyms=OSUNBa0004G10.24;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

RP SEQUENCE FROM N.A.
 RX PubMed=12447438; DOI=10.1038/nature01184;
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 Wu J., Nishimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 Hijikata S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuuchi A., Kamiya K.,
 Kikawa W., Katagiri S., Kikuta A., Kobayashi T., Kono I.,
 Machida K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 Nagaoka H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 Namiki N., Negishi M., Ono I., Ono N., Saji S., Sakai K., Shibata M.,
 Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 Zhang H., Iwama H., Bodo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 Yano M., Jiang J., Gojohori T.,
 RA "The genome sequence and structure of rice chromosome 1.",
 RT Nature 420:312-316(2002).
 RL PubMed=12447438; DOI=10.1038/nature01184;
 DR EMBL; AP003074; BAB18317.1; -;
 DR EMBL; AP003074; BAB40077.1; -;
 DR Gramene; Q9FTL5; -;
 SQ SEQUENCE 243 AA; 25227 MW; CDAEDSD0F5E2E CRC64;

Alignment Scores:
 Pred. No.: 58.5 Length: 243
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q9FTL5 (1-243)

QY 1064 TGGAGAGGAGGAGGCGCGCCCA 1090
 Db 74 TTPARGATGSLGSLGSLGSLAATG 82

RESULT 27
 Q73V06 PRELIMINARY; PRT; 274 AA.
 AC Q73V06; 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Nucl.
 GN Name=nucl; OrderedLocNames=MAP3210;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriinae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxId=1770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.,
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017238; AAS05758.1; -;
 DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
 DR InterPro; IPR001457; Oxidored_g3;
 DR Pfam; PF00499; Oxidored_g3; 1.
 KW Complete proteome.
 SQ SEQUENCE 274 AA; 28449 MW; B4A3B730ADFB6F68 CRC64;

Alignment Scores:
 Pred. No.: 57.6 Length: 274
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q73V06 (1-274)

QY 494 AGAGAGAGAGGAGTACGCGCGCC 520
 Db 115 ArgGlyGlnArgValAlaValAla 123

RESULT 28
 Q63SY8 PRELIMINARY; PRT; 284 AA.
 AC Q63SY8; 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative phage-related DNA polymerase.
 GN ORFNames=BPSL2183;
 OS Burkholderia pseudomallei K96243.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 OX NCBI_TaxId=272560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K96243;
 RX PubMed=15377794;
 RA Holden M.T.G., Tidball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
 Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
 Bentley S.D., Sebahia M., Thomson N.R., Bacon N., Beacham I.R.,
 Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
 Chillingworth T., Cronin A., Crosset B., Davis P., Deshaizer D.,
 Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagers K.,
 Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 Rabinowitz B., Rutherford K., Sanders M., Simmonds M.,
 Songsevivai S., Stevens C., Tumapa S., Vesaratchavee M.,
 Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.,
 RT "Genomic plasticity of the causative agent of melioidosis,
 RT Burkholderia pseudomallei".
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
 DR EMBL; BX571965; CAH36185.1; -;
 SQ SEQUENCE 284 AA; 30161 MW; 491795AFE38AA046 CRC64;

Alignment Scores:
 Pred. No.: 57.3 Length: 284
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q63SY8 (1-284)

QY 817 ACAAGTCAGAGGAGGCGCGCGCCT 791
 Db 40 ThrArgProAlaAlaAlaProAlaPro 48

RESULT 29
 Q855F3 PRELIMINARY; PRT; 287 AA.
 AC Q855F3; 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Gp57.
 OS Mycobacteriophage Cheb.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
 OX NCBI_TaxId=205868;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karchikyan T., Wadsworth C.,
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Panunzio N.R.,
 RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
 RA Hatfull G.F.,
 RT "Origins of highly mosaic mycobacteriophage genomes."
 RL Cell 113:171-182(2003).
 DR EMBL; AY129330; AAN12455.1; -;

DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR PRINTS; PR01590; HTHFIS.
SQ SEQUENCE 287 AA; 28977 MW; 9AFA47531C54D08B CRC64;

Alignment Scores:

Pred. No.:	57.2	Length:	287
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.45%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x Q855F3 (1-287)

QY 338 TGGAGCGCGCGGCGAGTGGCGGAGG 364
Db 158 TTPSerAlaIaGlySerAlaIaArg 166

RESULT 30
Q6YTS7
ID 06YTS7 PRELIMINARY; PRT; 296 AA.
AC 06YTS7;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DE Hypoetical protein P0419H09.23.
GN Name=P0419H09.23;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (NCV-2002) to the EMBL/GenBank/DBJ databases.
KW Hypoetical protein.
KW SEQUENCE 296 AA; 31159 MW; 37BC933E4B3E4A1A CRC64;

Alignment Scores:

Pred. No.:	57	Length:	296
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.45%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x Q6YTS7 (1-296)

QY 1652 TCACCTTCCCTCCGCTGCTCAGAGCT 1678
Db 29 SerProSerProLeuSerAlaGala 37

RESULT 31
Q6RYL0
ID 06RYL0 PRELIMINARY; PRT; 307 AA.
AC 06RYL0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE Hypoetical protein DRA0302.
GN OrderedLocustNames=DRA0302;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=RI / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Uitterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Manton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Frazer C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577 (1999).
DR EMBL; AE001863; AAF12516.1; -.
DR PIR; C75584; C75584.
DR TIGR; DRA0302; -.
KW Complete proteome; Hypoetical protein.
SQ SEQUENCE 307 AA; 32062 MW; EED6754B565DC8B CRC64;

Alignment Scores:

Pred. No.:	56.7	Length:	307
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x Q84Z99 (1-307)

QY 349 CGCGGCGCTCCAGCTCCAGCAGAGCT 323
Db 278 ArgGlyAlaProAlaProAlaGlyAla 286

RESULT 32
Q84Z99
ID 084Z99 PRELIMINARY; PRT; 310 AA.
AC 084Z99;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hypoetical protein P0705A05.130 (Hypoetical protein
DE OSUNBA0044E16.14).
GN Name=P0705A05.130; Synonyms=OSUNBA0044E16.14;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004623; BAC56007.1; -.
DR EMBL; AP005411; BADI0328.1; -.
DR Gramene; Q84Z99; -.
DR InterPro; IPR002345; Lipocalin.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
KW Hypoetical protein.
SQ SEQUENCE 310 AA; 32379 MW; 0E9B42DBF64B08A1 CRC64;

Alignment Scores:

Pred. No.:	56.7	Length:	310
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.45%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x Q84Z99 (1-310)

QY 1056 GTTGCTGATGAGAGGAGGAGGCG 1082
|||||

```

Db      188 ValAlaApGcLyGlyArgGlyAla 196
RESULT 33
ID      089J02      PRELIMINARY;      PRT;      335 AA.
AC      089J02;
DT      01-JUN-2003 (TEMBLrel. 24, Created)
DT      01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DE      B15482 protein.
GN      OrderedLocustNames=b115482;
OS      Bradyrhizobium japonicum.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Bradyrhizobium.
OX      NCBI_TaxID=375;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=USD110;
RX      MEDLINE=22484998; PubMed=12597275;
RA      Kaneo T., Nakamura Y., Sato S., Minamisawa K., Uchiyumi T.,
RA      Sasagawa S., Matsumoto M., Idehara K., Iriuchi M., Kawashima K.,
RA      Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA      Tabata S.;
RT      "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT      Bradyrhizobium japonicum USD110."
RL      DNA Res. 9:189-197(2002).
DR      EMBL: AP005955; BACS0747.1; -.
KM      Complete proteome.
SQ      SEQUENCE 335 AA; 37415 MW; B96871BC36F0953B CRC64;

Alignment Scores:
Pred. No.:      56.1      Length:      335
Score:      9.00      Matches:      9
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      1.45%      Indels:      0
DB:      Gaps:      0

US-10-030-271-3 (1-1883) x Q89J02 (1-335)
QY      85 GCGCGGGTCTGAGCTGTTCGCCGCTC 111
Db      278 GlyArgValLeuSerLeuPheArgLeu 286

RESULT 34
ID      08R353      PRELIMINARY;      PRT;      347 AA.
AC      08R353;
DT      01-JUN-2002 (TEMBLrel. 21, Created)
DT      01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DE      01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE      Thr3p3 protein.
GN      Mus musculus (Mouse).
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heile F.,
RA      Diatchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Udén T.B., Tosiyuki S., Carninci P., Prange C.T.,
RA      Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullan S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Rischard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

```

```

RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA      Jones S.J., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA      Straube R.;
RL      Submitted (Apr-2002) to the EMBL/GenBank/DBU databases.
DR      EMBL: BC026580; AAH26580.1; -.
DR      MGI: 2442637; Thr3p3.
DR      GO: GO:0000119; C-mediator complex; IDA.
DR      GO: GO:0006366; P:transcription from Pol II promoter; IC.
SQ      SEQUENCE 347 AA; 38293 MW; C885A3C2394FAD46 CRC64;

Alignment Scores:
Pred. No.:      55.8      Length:      347
Score:      9.00      Matches:      9
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      1.46%      Indels:      0
DB:      Gaps:      0

US-10-030-271-3 (1-1883) x Q8R353 (1-347)
QY      347 CGCGGCTCCAGCTCCAGCAGAGACTCT 321
Db      155 ArgArgSerSerSerArgSerSer 163

RESULT 35
ID      WNT1_HUMAN      STANDARD;      PRT;      370 AA.
AC      P04628;
DT      13-AUG-1987 (Rel. 05, Created)
DT      13-AUG-1987 (Rel. 05, Last sequence update)
DT      05-JUL-2004 (Rel. 44, Last annotation update)
DE      Wnt-1 proto-oncogene protein precursor.
GN      Name=WNT1; Synonyms=INT1;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=86055728; PubMed=2998762;
RA      van Ooyen A., Kwee V., Nusse R.;
RT      "The nucleotide sequence of the human int-1 mammary oncogene;
RT      evolutionary conservation of coding and non-coding sequences.";
RL      EMBO J. 4:2905-2909(1985).
CC      -1- FUNCTION: Ligand for members of the frizzled family of seven
CC      transmembrane receptors. Probable developmental protein. May be a
CC      signaling molecule important in CNS development. Is likely to
CC      signal over only few cell diameters.
CC      -1- SUBCELLULAR LOCATION: Possibly secreted and associates with the
CC      extracellular matrix.
CC      -1- SIMILARITY: Belongs to the Wnt family.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X03072; CYA26874.1; -.
DR      PIR: A24674; TVHUT1.
DR      Genew: HGNC:12774; WNT1.
DR      MIM: 164820; -.

```

DR GO; GO:0007417; P:central nervous system development; TAS.
 DR GO; GO:0005653; P:morphogenesis; TAS.
 DR InterPro; IPR005817; Wnt.
 DR InterPro; IPR009139; Wnt1.
 DR InterPro; IPR005816; Wnt_gnrfactor.
 DR Pfam; PF00110; wnt; 1.
 DR PRINTS; PRO184; WNT1PROTEIN.
 DR PRINTS; PRO1349; WNTPROTEIN.
 DR SMART; SM00097; WNT1; 1.
 DR PROSITE; PS00246; WNT1; 1.
 DR Developmental protein; Glycoprotein; Proto-oncogene; Signal;
 Wnt signaling pathway.
 FT SIGNAL 1 27 Potential.
 FT CHAIN 28 370 Wnt-1 proto-oncogene protein.
 FT CARBOHYD 29 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 316 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 346 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 359 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 359 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 370 AA; 40981 MW; F7EB11DA12E173F CRC64;

Alignment Scores:

Pred. No.: 55.4 Length: 370
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x WNT1_HUMAN (1-370)

QY 366 GCTCTGCGCCGACTGCGCGGGCGCTC 340

DB 18 AAlaLeuAlaLeuProAlaAlaLeu 26

RESULT 36

Q7Z5U1

PRELIMINARY; PRT; 373 AA.

ID Q7Z5U1 PRELIMINARY; PRT; 373 AA.
 AC Q7Z5U1;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE THRAP3 protein (Fragment).
 GN Name=THRAP3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rodriguez A.C., Grimwood J., Green E.D., Dickinson M.C.,
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Bosak S.A., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schmech A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.

RC TISSUE=Skin;
 RA Strausberg R.L.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC054046; AAH54046.1; -.
 FT NON TER 373 373
 SQ SEQUENCE 373 AA; 41600 MW; 05EA5A7DCFF97F9B CRC64;

Alignment Scores:

Pred. No.: 55.3 Length: 373
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q7Z5U1 (1-373)

QY 347 CGGCGCTGCGCTCCGAGGAGCTCT 321

DB 155 ArgArgSerSerSerArgSerSer 163

RESULT 37

Q6P0P7

PRELIMINARY; PRT; 374 AA.

ID Q6P0P7 PRELIMINARY; PRT; 374 AA.
 AC Q6P0P7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE THRAP3 protein (Fragment).
 GN Name=THRAP3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schmech A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC065519; AAH65519.1; -.
 FT NON TER 374 374
 SQ SEQUENCE 374 AA; 41729 MW; AF05EA5A75BFF97F CRC64;

Alignment Scores:

Pred. No.: 55.3 Length: 374
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q6P0P7 (1-374)

QY 347 GCGGCGCTCCAGCTCCAGCAGAGCTCT 321

DB 155 ArgArgSerSerSerSerArgSerSer 163

RESULT 38

Q6PJV4 PRELIMINARY; PRT; 375 AA.

AC Q6PJV4; 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DE THRAP3 protein (Fragment).

GN Name=THRAP3;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxId=9606;

[1]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta, and Skin;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,

RA Diatchenko L., Marisla K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Rahe S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,

RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunarrane P.H.,

RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzyzanski M.I., Skalska U., Small D.E., Schmechel A., Schein J.E.,

RA Jones S.J., Matra M.A.,

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[2]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA Strausberg R.L.

RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RA Strausberg R.L.

RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC010997; AA010997.1; -.

FT NON TER 375 375

SEQUENCE 375 AA; 41857 MW; 2B6F05EA5A75EFP9 CRC64;

Alignment Scores:

Pred. No.: 55.3 Length: 375

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.46% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q6PJV4 (1-375)

QY 347 GCGGCGCTCCAGCTCCAGCAGAGCTCT 321

DB 155 ArgArgSerSerSerSerArgSerSer 163

RESULT 39

Q82B67 PRELIMINARY; PRT; 382 AA.

AC Q82B67; 01-JUN-2003 (TREMBlrel. 24, Created)

DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)

DE Hypothetical protein.

GN OrderedLocusNames=SAV5838;

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomycetes.

NCBI_TaxId=33903;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680;

RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

RA Sakaki Y., Hattori M., Omura S.,

RT "Complete genome sequence and comparative analysis of the industrial

microorganism Streptomyces avermitilis."

Nat. Biotechnol. 21:526-531(2003).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RA Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,

RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,

RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,

RT "Genome sequence of an industrial microorganism Streptomyces

RT avermitilis: deducing the ability of producing secondary

RT metabolites."

Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

DR EMBL; AF005044; BAC73550.1; -.

SEQUENCE 382 AA; 41611 MW; FBC5D6358F1FAD7 CRC64;

Alignment Scores:

Pred. No.: 55.1 Length: 382

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.45% Indels: 0

DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q82B67 (1-382)

QY 783 GCGGCGCTCCAGCTCCAGCAGAGCTCT 809

DB 133 AlaAlaProGlyAlaGlyAlaAlaAla 141

RESULT 40

Q6M204 PRELIMINARY; PRT; 385 AA.

AC Q6M204; 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DE BETA-N-ACETYLGLUCOSAMINIDASE (BC 3.2.1.52).

GN Name=nagA2; OrderedLocusNames=cg3158;

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.

NCBI_TaxId=1718;

[1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

RA Kalinowski J., Bathe B., Barthele D., Bischoff N., Bock M.,

RA Burkowski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,

RA Goemann A., Hartmann M., Hutmacher K., Kraemer R., Linke B.,

RA McHardy A.C., Meyer F., Moeckel B., Pfefferle W., Puhler A.,

RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins.";
 RL J. Biotechnol. 104:5-25(2003).
 DR EMBL: BX927156; CAF20877.1; -.
 DR GO: GO:0004563; F-beta-N-acetylhexosaminidase activity; IEA.
 DR GO: GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . .; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro: IPR001764; Glyco_hydro_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 385 AA; 40473 MW; 367F80C810D53849 CRC64;

Alignment Scores:
 Pred. No.: 55.1 Length: 385
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q6W204 (1-385)

QY 967 GAGGCCCTGCGAGGCTGTGGCCGG 993
 DB 107 GUAAlaLeuArgGUAAlaValGlyArg 115

RESULT 41
 Q7XJW1 PRELIMINARY; PRT; 393 AA.

AC Q7XJW1;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE OSUNB0016002.9 protein.
 GN Name=OSUNB0016002.9;
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 OX [1]
 RN SEQUENCE FROM N.A.
 RX PubMed=12447439; DOI=10.1038/nature01183;
 RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
 RA Meng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
 RA Chen J., Ren S., Shao X., Shao X., Sun Y., Hu Q., Zhang X., Zhang W.,
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
 RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
 RA Han B.;
 RT "Sequence and analysis of rice chromosome 4.";
 RL Nature 420:316-320(2002).
 DR EMBL: AL606588; CAE05999.3; -.
 DR HSSP: P00909; IJCM.
 DR Gramene; Q7XJW1;
 DR GO: GO:0004425; F:tri-ole-3-glycerol-phosphate synthase activity; IEA.
 DR GO: GO:0006568; P:trypicophan metabolism; IEA.
 DR InterPro: IPR003009; FMN enzyme.
 DR InterPro: IPR001468; IGPS.
 DR InterPro: IPR01060; Ribp_bind_barrel.
 DR Pfam: PF00218; IGPS; 1.
 DR Pfam: PF001511; IGPS; 1.
 DR PROSITE: PS00614; IGPS; 1.
 SQ SEQUENCE 393 AA; 42409 MW; 80B9F1860BDE8547 CRC64;

Alignment Scores:
 Pred. No.: 54.9 Length: 393

Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q7XJW1 (1-393)

QY 424 GCGGACGAGCTGTGGCGGCGACCA 398
 DB 6 AlaAlaAlaGlyArgGlyGlyProAla 14

RESULT 42
 Q8NLT5 PRELIMINARY; PRT; 395 AA.

AC Q8NLT5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Beta-glucosidase-related glycosidases (EC 3.2.1.30).
 GN OrderedLocusNames=Cg12852;
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OC NCBI_TaxID=1718;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
 RA Nakagawa S.;
 RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP005283; BAC00246.1; -.
 DR GO: GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . .; IEA.
 DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro: IPR001764; Glyco_hydro_3N.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 KW Complete proteome; Glycosidase; Hydrolase.
 SQ SEQUENCE 395 AA; 41554 MW; FB35F878C4F06AD CRC64;

Alignment Scores:
 Pred. No.: 54.9 Length: 395
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q8NLT5 (1-395)

QY 967 GAGGCCCTGCGAGGCTGTGGCCGG 993
 DB 117 GUAAlaLeuArgGUAAlaValGlyArg 125

RESULT 43
 Q7MMW4 PRELIMINARY; PRT; 410 AA.

AC Q7MMW4;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative integral membrane protein.
 GN OrderedLocusNames=BB1276;
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OC NCBI_TaxID=516;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
 RA Parkhill J., Sebatina M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,

RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabbittowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.",
RL Nat. Genet. 35:32-40(2003).
DR EMBL: BX640440; CAB31774.1; -.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005215; F: transporter activity; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR007114; MFS.
DR PROSITE: PS50850; MFS; 1.
KM Complete proteome.
SQ SEQUENCE 410 AA; 40657 MW; 135C99999DC40AF0 CRC64;

Alignment Scores:
Pred. No.: 54.6 Length: 410
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q7MMW4 (1-410)
QY 780 CCGGCGGCGCCGCGCTGCGCGCGCA 806
DB 10 ProAlaAlaProGlyAlaGlyAlaAla 18

RESULT 44
Q7Q5N2 PRELIMINARY; PRT; 447 AA.
ID Q7Q5N2
AC Q7Q5N2;
DT 01-MAR-2004 (TEMBLrel. 26, Created)
DT 01-MAR-2004 (TEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE AgcP6172 (Fragment).
GN Name=agcG50969; ORFName=ENSANG0000010497;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008960; EAA10851.1; -.
DR GO: GO:0016020; C: membrane; IEA.
DR GO: GO:0005542; F: folate acid binding; IEA.
DR GO: GO:0008518; F: reduced folate carrier activity; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR002666; Folate_carrier.
DR Pfam: PF01770; Folate_carrier; 1.
DR TIGRFAMs: TIGR00806; ffc; 1.
FT NON_TER 1
FT SEQUENCE 447 AA; 49943 MW; BBB75507C29302CF CRC64;

Alignment Scores:
Pred. No.: 54 Length: 447
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q7Q5N2 (1-447)
QY 1700 ATCCAGTGCAAGTCATGTTGACGTC 1674
DB 245 IleglnvalGlnvalGlnvalGlnleu 253

RESULT 45
Q7VUX6 PRELIMINARY; PRT; 452 AA.
ID Q7VUX6
AC Q7VUX6;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Putative integral membrane protein.
GN OrderedLocustNames=BP2940;
OS *Bordetella pertussis*.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; *Bordetella*.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabbittowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.,
RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.",
RL Nat. Genet. 35:32-40(2003).
DR EMBL: BX640419; CAB43212.1; -.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0005215; F: transporter activity; IEA.
DR GO: GO:0006810; P: transport; IEA.
DR InterPro: IPR007114; MFS.
DR PROSITE: PS50850; MFS; 1.
KM Complete proteome.
SQ SEQUENCE 452 AA; 45805 MW; A369C2AC3B174191 CRC64;

Alignment Scores:
Pred. No.: 54 Length: 452
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q7VUX6 (1-452)
QY 780 CCGGCGGCGCCGCGCTGCGCGCGCA 806
DB 54 ProAlaAlaProGlyAlaGlyAlaAla 62

RESULT 46
Q8YXA3 PRELIMINARY; PRT; 454 AA.
ID Q8YXA3
AC Q8YXA3;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Alr1312 protein.
GN OrderedLocustNames=alr1312;
OS *Anabaena* sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840; Kuritz T., Saeamoto S.,
 RA Kaneko T., Nakamura Y., Wolk C.P.,
 RA Watanabe A., Iritsuuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsuno M., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.,
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003585; BAB73269.1; -
 DR PIR: AE1970; AE1970.
 DR GO: GO:0003824; F: catalytic activity; IEA.
 DR GO: GO:0005056; F: iron ion binding; IEA.
 DR Pfam: PF04055; Radical_SAM; 1.
 DR Pfam: PF01938; TRAM; 1.
 DR Pfam: PF00919; UPF0004; 1.
 DR SMART: SM00729; ELP3; 1.
 DR TIGRFAMs: TIGR01574; mab-methiolase; 1.
 DR TIGRFAMs: TIGR00089; UPF0004; 1.
 DR PROSITE: PS50926; TRAM; 1.
 DR PROSITE: PS01278; UPF0004; 1.
 KM Complete proteome.
 SQ SEQUENCE 454 AA; 51509 MW; 8606FD00B3D7BDBC CRC64;
 Alignment Scores:
 Pred. No.: 53.9 Length: 454
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 2 Gaps: 0
 US-10-030-271-3 (1-1883) x Q8BV76 (1-454)
 QY 651 AGCGGACCCGACGACGTCAGACGC 677
 Db 175 SerrArgThrProAlaAlaValArgAla 183
 RESULT 47
 Q8BV76 PRELIMINARY; PRT; 458 AA.
 AC Q8BV76;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus adult male spinal cord cDNA, RIKEN full-length enriched
 DE library clone: A33062C13 product: similar to THYROID HORMONE RECEPTOR-
 DE ASSOCIATED PROTEIN COMPLEX COMPONENT TRAP150 (Fragment).
 GN Name=Thrap3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.,
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RA the FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe K.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:11757-1771(2000).
 RN [6]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Spinal cord;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohmoto N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK079601; BC037696.1; -
 DR MGI: 2442637; Thrap3.
 DR GO: GO:0000119; C: mediator complex; IDA.
 DR GO: GO:0006366; P: transcription from Pol II promoter; IC.
 KM Receptor.
 FT NON TER 458 458
 SQ SEQUENCE 458 AA; 51125 MW; 63E03903164CB1ED CRC64;
 Alignment Scores:
 Pred. No.: 53.9 Length: 458
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 2 Gaps: 0
 US-10-030-271-3 (1-1883) x Q8BV76 (1-458)
 QY 347 CGGCGCTCCAGCTCCGACGACGCTCT 321
 Db 155 ArgArgSerSerSerArgSerSer 163
 RESULT 48
 FUNC_SYNXP STANDARD; PRT; 465 AA.
 AC Q7U8I1;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Fumarate hydratase class II (EC 4.2.1.2) (Fumarase C).
 GN Name=func; OrderedLocusNames=STYM06377;
 OS Synecococcus sp. (strain WH8102).


```

OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN [1]
RX MEDLINE=2825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahmsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerlin J.E., Regala W., Allen E.E., McCareen J.,
RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus";
RL Nature 424:1037-1042(2003).
CC -1- CATALYTIC ACTIVITY: (S)-malate = fumarate + H(2)O.
CC -1- PATHWAY: Tricarboxylic acid cycle.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the class-II fumarase / aspartase family.
Fumarase subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BX569690; CAB07152.1; ALT_INIT.
DR HAMAP; MF_00743; -. 1.
DR InterPro; IPR003031; D_crysalin.
DR InterPro; IPR000362; Fumarate_lyase.
DR Pfam; PF00206; Lyase_1; 1.
DR PRINTS; PR00145; DCRYSTALIN.
DR PROSITE; PS00149; FUMARATE_LYASES.
DR PROSITE; PS00163; FUMARATE_LYASES; 1.
KM Complete proteome; Lyase; Tricarboxylic acid cycle.
FT ACT_SITE 189 189 Potential.
FT BINDING 325 325 Substrate carboxyl (Potential).
SQ SEQUENCE 465 AA; 50057 MW; CA673J3D8C464D CRC64;

Alignment Scores:
Pred. No.: 53.8 Length: 465
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 1 Gaps: 0

US-10-030-271-3 (1-1883) x FUMC_SYNPK (1-465)
QY 348 GCGGCGCTCCAGCTCCAGAGAGCTC 322
Db 153 AlaAlaLeuGlnLeuGlnGlnGlnLeu 161

RESULT 49
Q910P9 PRELIMINARY; PRT; 468 AA.
ID Q910P9;
AC Q910P9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative membrane protein.
GN ORFNames=SCDBA.15c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RX STRAIN=MA-4680;
RA MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleeer H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornby T., Howarth S.,

```

```

RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neil S.,
RA Rabinowitsch B., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
DR EMBL; AL939119; CAB77336.1; -.
DR InterPro; IPR008979; Gal_bind_like.
KM Complete proteome.
SQ SEQUENCE 468 AA; 48544 MW; C0FD487AA31CAC0 CRC64;

Alignment Scores:
Pred. No.: 53.7 Length: 468
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x Q910P9 (1-468)
QY 783 GCGGCGCTCCAGCTCCAGAGAGCTC 809
Db 118 AlaAlaProGlyAlaGlyAlaAlaAla 126

RESULT 50
Q82HT1 PRELIMINARY; PRT; 478 AA.
ID Q82HT1;
AC Q82HT1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative GntR-family transcriptional regulator.
GN OrderedLocNames=SAV3427;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=33903;
RN [1]
RX STRAIN=MA-4680;
RA MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RX STRAIN=MA-4680;
RA MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- SIMILARITY: Contains 1 HTH gntR-type DNA-binding domain.
DR EMBL; AP005035; BAC71139.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000524; HTH_GNTR.
DR InterPro; IPR000408; Reg_chr_condens.
DR InterPro; IPR009058; Wing_chr_bnd.
DR Pfam; PF00392; GntR_1.
DR PRINTS; PR00035; HTHGNTR.
DR SMART; SM00345; HTH_GNTR. 1.
DR PROSITE; PSS0949; HTH_GNTR. 1.
DR PROSITE; PSS0626; RCG1_2; UNKNOWN_1.

```

KW Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 478 AA; 50879 MW; 917EB2C508E3496E CRC64;

Alignment Scores:

Pred. No.:	53.6	Length:	478
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.45%	Indels:	0
DB:	2	Gaps:	0

US-10-030-271-3 (1-1883) x Q82HT1 (1-478)

QY	961	CTGACTGAGGCCCTGCCAGAGGCTGTG	987
Db	25	LeuThrGluAlaLeuArgGluAlaVal	33

Search completed: March 22, 2005, 03:20:51
 Job time : 360.771 secs


```
84      8      1.3      10      4      AAG95334      Human com
85      8      1.3      11      6      AAO31170      Pilon (Pr
86      8      1.3      15      2      AAW37662      Peppy moc
87      8      1.3      15      2      AAW38911      Peptide r
88      8      1.3      15      7      ADB49288      Novel MW
89      8      1.3      21      4      AAB45725      Human 7TM
90      8      1.3      32      8      ADO05533      Matze (AS
91      8      1.3      45      5      ABP43177      Human ova
92      8      1.3      50      6      ABM46314      Propionib
93      8      1.3      50      6      ABM46314      Propionib
94      8      1.3      52      2      AAY20499      Human neu
95      8      1.3      60      4      ABG04482      Novel hum
96      8      1.3      62      8      ABO58857      Human gen
97      8      1.3      63      2      AAY24144      Cons. tex
98      8      1.3      65      7      ABO62516      Klebsiell
99      8      1.3      73      4      ABG00429      Novel hum
c 100      8      1.3      73      4      ABG13593      Novel hum
```

ALIGNMENTS

```
RESULT 1
AAB60387
ID AAB60387 standard; protein: 326 AA.
```

```
AC AAB60387;
```

```
DT 24-APR-2001 (first entry)
```

```
DE Human apoptosis-associated factor NT2RM1000558, SEQ ID NO:4.
```

```
KW Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
```

```
KM DED; caspase family cleavage domain; pro-apoptotic; drug screening;
```

```
KM cell proliferation; ischaemic disease; chronic viral disease.
```

```
XX Homo sapiens.
```

```
OS Homo sapiens.
```

```
PN WO200104300-A1.
```

```
XX 18-JAN-2001.
```

```
PF 06-JUL-2000; 2000MO-JP004516.
```

```
PR 08-JUL-1999; 99JP-00194179.
```

```
PR 18-OCT-1999; 99US-0159586P.
```

```
XX (HELI-) HELIX RES INST.
```

```
PI Oca T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
```

```
XX WPI; 2001-138348/14.
```

```
DR N-PSDB; AAF27408.
```

```
PT Polynucleotide encoding an apoptosis-associated factor protein with death
```

```
PT effector domain and caspase family-cleavage domain, useful in regulating
```

```
PT diseases with cell proliferation.
```

```
XX Claim 3; Page 47-48; 53pp; Japanese.
```

```
XX The invention relates to a novel human apoptosis-associated factor
```

```
CC (AAB60387), designated NT2RM1000558, which contains a death
```

```
CC effector domain (DED) and a caspase family cleavage domain and is capable
```

```
CC of inducing apoptosis in cells. The invention also relates to nucleic
```

```
CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
```

```
CC (particularly dominant negative variants); vectors and host cells
```

```
CC comprising a nucleic acid which encodes an apoptosis-associated factor
```

```
CC of the invention; the recombinant production of the protein; an antibody
```

```
CC against the protein; and methods of screening for compounds which can
```

```
CC regulate apoptosis. The apoptosis-related factor is useful in regulating
```

```
CC diseases associated with cell proliferation and in screening drug
```

```
CC candidates e.g., for regulating cell proliferation or cell death in
```

```
CC ischaemic diseases and chronic viral diseases. The present sequence
```

```
CC represents the human apoptosis-associated factor NT2RM1000558
```

```
XX      SQ      Sequence 326 AA:
Alignment Scores:
Pred. No.: 1,53e-295
Score: 326.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 52.58%
DB: 4
Gaps: 0
```

```
US-10-030-271-3 (1-1883) x AAB60387 (1-326)
```

```
QY 124 ATGGCCGCTATCCGGGCTGACCCCGCCCGTCTGGAGAGAGATGATGCTTCACTAC 183
```

```
Db 1 MetAlaLeuSerGlySerThrProAlaProCyStrpGluGluAbpGluCySerLeuApTyx 20
```

```
QY 184 TACGGGATGCTGCTGCTTCAACCGTATGCTGAGTGGTGGGCGGCACTGACGAGTGC 243
```

```
Db 21 TyrGlyMetLeuSerLeuHisArgMetPheGluValValGlyGluLeuThrGluCyS 40
```

```
QY 244 GAGCTGAGGCTCTGGCTTTCTGTGTGATGAGGCTCTGGCCCGCGAGGCTTAGCC 303
```

```
Db 41 GluLeuGluLeuLeuAlaPheLeuLeuLeuAlaProGlyAlaAlaGlyLeuAla 60
```

```
QY 304 CGGCGCCGCGACCGGCTTAGGCTCTGCTGAGCTGAGCGCGCGGCGAGTGGCGAG 363
```

```
Db 61 ArgAlaArgSerGlyLeuGluLeuLeuGluLeuGluLeuArgGlyGlyGlyGlu 80
```

```
QY 364 AGCAACCTGCGGCTGCTGGGCGCAACTCTGCGCGCTGCTGGCCCGCGACGACTGTCGC 423
```

```
Db 81 SerAsnLeuArgLeuLeuGlyGluLeuLeuValLeuAlaArgHisAspLeuLeuPro 100
```

```
QY 424 CACCTGCGCGCAGCGCGCGCCGCGCAGTGTCTCCAGAACGCTTAGTGGACCTCC 483
```

```
Db 101 HisLeuAlaArgIysArgArgArgProValSerProGluArgGlySerTyxGlyThrSer 120
```

```
QY 484 AGCTCTTCAAAAGACACAGAGGGTATGCTGCTGCGCGCTGCGGACGTCAGCACTTGC 543
```

```
Db 121 SerSerSerIysArgThrGluGlySerCyArgArgArgGlnSerSerSerAla 140
```

```
QY 544 AATTCTCAGAGGGTCACTGCGGAGACAGGCTCCCGCCCAACCAAGGGCGCGCGGAGT 603
```

```
Db 141 AsnSerGlnGlnGlyGlnTrpGlnTrpGlySerProProThrIysArgGlnArgSer 160
```

```
QY 604 CGGCGCGCGCCAGTGGTGTGCGCAGACGCGCGCGAGAGGGGCCCGCGACGCCAG 663
```

```
Db 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
```

```
QY 664 CAGCAGTCAGAGCCCGCGCAGACCTTCTCTGAAGGCAAGTGAATCTGTGATCCGGCTC 723
```

```
Db 181 GlnGlnSerGluProAlaArgProSerSerGlnGlyValValThrCyAspIleArgLeu 200
```

```
QY 724 CGGGTTCGACAGAGTACATGCGGAGCATGGGCGAGGCTGGAGGAGCGGCGATCCCGG 783
```

```
Db 201 ArgValArgAlaGluTyrCySglnHisGlyProAlaLeuGlnGlnGlyValAlaSerArg 220
```

```
QY 784 CGGCGCCGAGCGCTGGCGCGCGCAGCTGACGCTGTTTGGGAGGCCACCGCAGTGTCCGC 843
```

```
Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
```

```
QY 844 TCAAGGACCTGGAGCTGTGTGTTTGTGACATCAAGTTCTCAAGCTCTCTATCTGGAC 903
```

```
Db 241 SerArgspLeuGlySerValValCyAspIleLysPheSerGlnLeuSerTyxLeuAsp 260
```

```
QY 904 GCCTTCGGGCGACATCAACGAGTGGCGGCTGCTGAGGCGCTGCGGGCGGCTTCTCG 963
```

```
Db 261 AlaPheThrPleIysAspTyxLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
```

```
QY 964 ACTGAGCCCTGCGAGAGGCTGTGGGCGCGGAGGCTGTGCTGTGCTGCTGCTGCTGCTGAT 1023
```

```
Db 281 ThrGluAlaLeuArgGluAlaValGlyArgGluAlaValAlaArgLeuLeuValSerValAsp 300
```

QY 1024 GAGGCTGACTATGAGCTGCGCCGCGCTGTGCTGATGAGAGGAAAGGGGGCGG 1083
 Db 301 GluAlaAspTyrGluAlaGlyArgArgLeuLeuMetGluGluGlyArg 320
 QY 1084 CGCCGACAGAGCGCTCC 1101
 Db 321 ArgProThrGluAlaSer 326
 RESULT 2
 AAB60386
 ID AAB60386 standard; protein; 303 AA.
 AC AAB60386;
 XX 24-APR-2001 (first entry)
 DE Human apoptosis-associated factor NT2RM1000558 (portion), SEQ ID NO:2.
 XX Human apoptosis-associated factor; NT2RM1000558; death effector domain;
 KM DED; caspase family cleavage domain; pro-apoptotic; drug screening;
 XX cell proliferation; ischaemic disease; chronic viral disease.
 OS Homo sapiens.
 PN WO200104300-A1.
 XX 18-JAN-2001.
 PD 06-JUL-2000; 2000MO-JP004516.
 XX 08-JUL-1999; 99JP-00194179.
 PR 18-OCT-1999; 99US-0159586P.
 XX (HELI-) HELIX RES INST.
 PA Oca T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
 PI MPI; 2001-138348/14.
 DR N-PSDB; AAF27407.
 XX Polynucleotide encoding an apoptosis-associated factor protein with death
 PT effector domain and caspase family-cleavage domain, useful in regulating
 PT diseases with cell proliferation.
 XX
 PS Claim 1; Page 43-44; 53pp; Japanese.
 XX
 CC The invention relates to a novel human apoptosis-associated factor
 CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
 CC effector domain (DED) and a caspase family cleavage domain and is capable
 CC of inducing apoptosis in cells. The invention also relates to nucleic
 CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
 CC (particularly dominant negative variants); vectors and host cells
 CC comprising a nucleic acid which encodes an apoptosis-associated factor
 CC of the invention; the recombinant production of the protein; an antibody
 CC against the protein; and methods of screening for compounds which can
 CC regulate apoptosis. The apoptosis-related factor is useful in regulating
 CC diseases associated with cell proliferation and in screening drug
 CC candidates e.g., for regulating cell proliferation or cell death in
 CC ischemic diseases and chronic viral diseases. The present sequence
 CC represents a substantial proportion of the human apoptosis-associated
 CC factor NT2RM1000558
 XX
 SQ Sequence 303 AA;
 Alignment Scores:
 Pred. No.: 5.1e-274 Length: 303
 Score: 303.00 Matches: 303
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.87% Indels: 0
 DB: 4 Gaps: 0
 US-10-030-271-3 (1-1883) x AAB60386 (1-303)

QY 124 ATGGCGTATCCGGGTGACACCCCGGCCCTGTGTGGAGAGATGATGCTGACTAC 183
 Db 1 MetAlaLeuSerGlySerThrProAlaProCybTirpGluLubspGluCybLeuAspTyr 20
 QY 184 TACGGATGCTGTGCTTACCCGTATGTTGAGAGTGTGGCGGCAACTGACCGACTGC 243
 Db 21 TyrGlyMetLeuSerLeuHisArgMet PheGluValValGlyGlyLeuThrGluCys 40
 QY 244 GAGCTGAGCTCTGGCTTTCTGTGATGAGAGCTCTGCGCCCGCGGAGGCTTAGCC 303
 Db 41 GluLeuGluLeuLeuAlaPheLeuLeuAspGluAlaProGlyAlaAlaGlyGlyLeuAla 60
 QY 304 CGGCGCCGACCGGCTTACAGCTCTGCTGAGAGCTGAGCGCGCGGCGAGTGGCGAG 363
 Db 61 ArgAlaArgSerGlyLeuGluLeuLeuLeuGluLeuGluLubspGlyGlyGlyGlyGly 80
 QY 81 SerAsnLeuArgLeuLeuGlyGlyGlyLeuLeuValLeuValAlaArgHisAspLeuLeuPro 100
 QY 424 CACCTGGCGGCAAGCGCGCGCGGCAAGTGTCCAGAAAGCTATGACAGCTCC 483
 Db 101 HisLeuAlaArgLysArgArgArgProValSerProGluArgTyrSerTyrGlyThrSer 120
 QY 484 AGCTTTCAAAAGAGACAGAGAGGTAGCTGCGCTGCGCGAGTCAAGTCTGCA 543
 Db 121 SerSerSerLysArgThrArgLysSerCybArgArgArgArgLubspSerSerSerAla 140
 QY 544 AATTTCAGAGGATGATGAGAGAGAGTCTCCCGGCAAGCGCGAGCGGAGT 603
 Db 141 AsnSerGlnGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 160
 QY 604 CGGCGCGCGGCGAGTGTGTGCGAGACCGCGCGGAGAGGCGCGAGCGGAG 663
 Db 161 ArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaAlaProGln 180
 QY 664 CACGAGTCAAGCGCGCGAGAGCTTCTGTAAGGCAAGTGAATGATCTGACCTCGGCTC 723
 Db 181 GlnGlnSerGluProAlaArgProSerSerGlnGlyValValThrCybAspLubspLeu 200
 QY 724 CGGCTTCAGACAGTATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 783
 Db 201 ArgValArgAlaGluTyrCybGluHisGlyProAlaLeuGlnGlnGlyValAlaSerArg 220
 QY 784 CGGCGCGGCGGCTGTGTGCGGCGGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 843
 Db 221 ArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeuArg 240
 QY 844 TCAAGGAGTGGGCTGTGTGTTTGATCATCAAGTTCAGAGCTCTCTATCTGGAC 903
 Db 241 SerArgAspLeuValValValValValValValValValValValValValValValVal 260
 QY 904 GCCTTCGGGCGACTACCTGAGTGGCGGCTGTGAGGCGCTGTGAGGCGGCTGTGCTG 963
 Db 261 AlaPheThrGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeu 280
 QY 964 ACTGAGGCTGTGAGAGGCTGTGGGCGGAGAGGCTGTGCTGTGCTGTGCTGTGAT 1023
 Db 281 ThrGluAlaLeuArgGluAlaValAlaGlyArgGluAlaValArgLeuLeuValSerValAsp 300
 QY 1024 GAGGCTGAC 1032
 Db 301 GluAlaAsp 303
 RESULT 3
 AAG73852
 ID AAG73852 standard; protein; 366 AA.
 XX AAG73852;
 AC AAG73852;
 XX 03-SEP-2001 (first entry)
 DT
 XX

DE Human colon cancer antigen protein SEQ ID NO:4616.
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma.
 XX
 OS Homo sapiens.
 XX WO200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX 28-SEP-2000; 2000WO-US026524.
 PF
 XX 29-SEP-1999; 99US-0157137P.
 PR 03-NOV-1999; 99US-0163280P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI WPI; 2001-235357/24.
 DR N-PSDB; AAH33283.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers.
 XX
 XX Claim 11; Page 6413-6414; 9803pp; English.
 PS
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patient's own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922
 CC
 CC
 CC Sequence 366 AA:
 SQ
 Alignment Scores:
 Pred. No.: 5,45e-231 Length: 366
 Score: 257.00 Matches: 357
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 1
 Query Match: 41.45% Gaps: 0
 DB: 4
 US-10-030-271-3 (1-1883) x AAG73852 (1-366)

Db 69 MetPheGluValAlaGlyGlyGlnLeuThrGluCysGluLeuGluLeuAlaPheLeu 88
 Qy 268 CTGGATGAGGCTCTCGGCGCCCGGAGGCTTAGACCGGCGGCGGCTAGAGCTC 327
 Db 89 LeuAspGluAlaProGlyAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGluLeu 108
 Qy 328 CTGCTGGAGCTGAGGCGCGGCGGCGAGTGGCGGAGAGCACTGGGCTGCGGGGCAA 387
 Db 109 LeuLeuGluLeuGluAlaArgGlyGlnCysAspGluSerAsnLeuArgLeuLeuGlyGln 128
 Qy 388 CTCCTGCGGCTGGCGCGGCGGCGAGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGG 447
 Db 129 LeuLeuArgValLeuAlaAlaArgHisAspLeuLeuProHisLeuAlaArgHisArgGly 148
 Qy 448 CCAAGTCTCCAGAACGCTATAGCTATAGCACTCCAGCTCTTCAAGAGAGAGAGGCT 507
 Db 149 ProValSerProGluArgGlySerGlyGlyThrSerSerSerSerSerSerSerSerSer 168
 Qy 508 AGCTGCGCTGCGGCTGGGAGTCAAGCACTTCTGCAATTTCTCAGCAGGGTCACTGGGAG 567
 Db 169 SerCysArgArgArgArgGlnSerSerSerSerSerSerSerSerSerSerSerSerSer 188
 Qy 568 ACAGGCTCCCGCCCAACCAAGCGGCGGCGGAGTCCGGGCGGCGGCGGCGGCGGCGGCGG 627
 Db 189 ThrGlySerProThrProThrArgGlnArgGlnArgSerArgGlyArgProSerGlyGlyAla 208
 Qy 628 AGACGGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 687
 Db 209 ArgArgArgArgArgGlyAlaProAlaAlaProGlnGlnGlnSerGlyProAlaArgPro 228
 Qy 688 TCCTCTGAAGGCAAAATGACCTGTGACATCCGCGCTCCGGGTTGAGCAGAGTACTGCGAG 747
 Db 229 SerSerGluGlyValThrCysAspIleArgLeuArgValArgAlaGluTyrCysGlu 248
 Qy 748 CATGGGCGGCGGCTGGAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 807
 Db 249 HisGlyProAlaLeuGlnGlnGlyValAlaSerArgArgProGlnAlaLeuAlaArgGln 268
 Qy 808 CTGAGCTGTTGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 867
 Db 269 LeuAspValPheGlyGlnAlaThrAlaValAlaLeuArgSerArgAspLeuGlySerVal 288
 Qy 868 TGTGACATCAAGTTTCAGAGCTCTCTATCTGGAAGCTTGGGGCGGAGTACCTCGAGT 927
 Db 289 CysAspIleLeuAspHisSerGluLeuSerTyrLeuAspAlaPheTrpGlyAspTyrLeuSer 308
 Qy 928 GGGCGGCTGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 987
 Db 309 GlyAlaLeuLeuGlnAlaLeuArgGlyValPheLeuThrGluAlaLeuArgGluAlaVal 328
 Qy 988 GGGCGGAGGCTGTTGCGCTGCTGTGATGTGATGAGGCTGATGAGGCTGAGCGCGG 1047
 Db 329 GlyArgGluAlaValAlaArgLeuLeuValSerValAspGluAlaAspTyrGluAlaArg 348
 Qy 1048 CGCGGCTGTTGCTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCT 1101
 Db 349 ArgArgLeuLeuLeuMetGluGluGlyGlyArgArgProThrGluAlaSer 366
 RESULT 4
 AAB15551
 ID AAB15551 standard; protein; 304 AA.
 XX
 XX AAB15551,
 AC
 XX
 XX 28-FEB-2001 (first entry)
 DT
 XX
 XX
 DE Apoptosis related protein encoded by gene 1 clone HXDOK36.
 XX
 XX Cytostatic; antiinflammatory; immunosuppressive; antisclerotic; cardiant;
 KW virucidal; anti-AIDS; vasotrophic; anti-ischaemic; antiParkinsonian;
 KW anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;
 KW colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;
 KW autoimmune disorder; multiple sclerosis; viral infection.

XX OS Homo sapiens.
 XX PN MO200056752-A2.
 XX PD 28-SEP-2000.
 XX PF 15-MAR-2000; 2000MO-US006642.
 XX PR 24-MAR-1999; 99US-0126018P.
 XX PR 17-JUN-1999; 99US-0139638P.
 XX PR 18-AUG-1999; 99US-0149449P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ruben SM, Ni J, Young PA;
 XX DR WPI; 2000-587660/55.
 XX DR N-PSDB; AAA95790.
 XX PT Nucleic acids encoding human apoptosis associated protein, useful for the
 PT prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's
 PT disease, inflammation and ischemic injury.
 PS Claim 1; Page 252-253; 273pp; English.
 XX CC The invention relates to the isolation of genes encoding 9 human
 CC apoptosis-related proteins. The nucleotide sequences AA95790-A95798
 CC encode the human apoptosis related proteins AAB15551-B15559. The genes
 CC can be used to generate fusion proteins by linking to the gene for the
 CC human immunoglobulin G Fc (IgG Fc) portion (AA95799) for increasing the
 CC stability of the fusion protein as compared to the human protein only.
 CC The gene and encoded protein may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate apoptosis associated
 CC protein expression, e.g. cancer (e.g. colon, breast and prostate cancer,
 CC melanomas and lymphomas), inflammation, autoimmune disorders (e.g.
 CC multiple sclerosis) and viral infections (e.g. herpes).
 XX SQ Sequence 304 AA;
 Alignment Scores:
 Pred. No.: 1,27e-221 Length: 304
 Score: 247.00 Matches: 247
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 39.84% Indels: 0
 DB: 3 Gaps: 0
 US-10-030-271-3 (1-1883) x AAB15551 (1-304)
 QY 361 GAGAGCAACTGGGGCTGTGGGGCAACTCTGTGGCGTGGCGCCGCAAGACTGTG 420
 DB 58 GUSerAmuLeuNrgLeuLeuGlyGlnLeuLeuNrgValLeuAlaArgHisAepLeuLeu 77
 QY 421 CCGCACCTGGCGCGCAAGCGCGCGCGCAAGTGTCTTCAGAACGCTATAGTGACACC 480
 DB 78 ProHieLeuAlaArgLysArgArgArgProValSerProGluArgLysSerTYrGlyThr 97
 QY 481 TCACAGCTCTTCAAGAGAGCAAGAGGGTAGTCCGCTGGCGCGCGCTCAAGCAGTTCT 540
 DB 98 SerSerSerSerLysArgLysArgLysArgLysSerCysArgArgArgArgGlnSerSerSer 117
 QY 541 GCAAAATTCAGAGCGGTGAGTGAGAGAGCGCTCCCGCCCAACCAAGCGCGCGCGCGG 600
 DB 118 AlaSerSerLysArgLysArgLysArgLysArgLysArgLysArgLysArgLysArgLys 137
 QY 601 AGTCGGGGCGCGCCAGTGTGTGTCCAGACGCGCGCGAGAGGGCGCCGACGCGCACCC 660
 DB 138 SerArgGlyArgProSerArgLysAlaArgArgArgArgArgGlyAlaProAlaAlaPro 157
 QY 661 CAGCAGCAGTCAGAGCGCGCCAGACCTTCTCTGAAGGCAAGTGACCTGTGACATCCGG 720
 DB 158 GlnGlnGlnSerGlnProAlaArgProSerSerSerLysValLysValThrCysAapLleArg 177

QY 721 CTCGGGTTTCAGACAGTACTCGAGCATGGCCAGACCTTTGGAGCAGGGCGTGGCATCC 780
 DB 178 LeuArgValArgLysArgLysArgLysArgLysArgLysArgLysArgLysArgLysArgLys 197
 QY 781 CGCGCGCGCCAGCGCGCTGGCGCGCGCGAGCTGGAGCTGTTGGGAGCGCACCGCAGTCTG 840
 DB 198 ArgArgProGlnAlaLeuAlaArgGlnLeuAlaSerValPheGlyGlnAlaThrAlaValLeu 217
 QY 841 CGCTCAAGGAGCCTGGGGCTGTGGTTTGATCATCAAGTTCAGAGCTCTCCTATCTG 900
 DB 218 ArgSerArgAspLeuGlySerValValCysAapLleLysSerGlnLeuSerTYrLeu 237
 QY 901 GACGCTTCTGGGCGCACTTACCTAGATGGCGCCCTGTGCGAGGCGCTTCGCGGCGTTC 960
 DB 238 AspAlaPheTrpGlyAspTYrLeuSerGlyAlaLeuLeuGlnAlaLeuArgLysValPhe 257
 QY 961 CTGACTGAGGCGCTTCGAGAGAGCTGTGGCGCGGAGGCTGTTGCGCTGTGTCAGTGTG 1020
 DB 258 LeuThrGlnAlaLeuAlaArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerVal 277
 QY 1021 GATGAGCTGACTATGAGGCTGGCGCGCGCTGTGCTGATGAGAGGAGGAGG 1080
 DB 278 AspGlnAlaAspTYrGlnAlaGlyArgArgArgLeuLeuLeuGlnGlnGly 297
 QY 1081 CGCGCGCGCCAGAGGCGCTTC 1101
 DB 298 ArgArgProThrGlnAlaSer 304
 RESULT 5
 ID ABB07263 standard; protein; 326 AA.
 ABB07263;
 DB 298 ArgArgProThrGlnAlaSer 304
 AC ABB07263;
 XX DT 26-MAR-2002 (first entry)
 XX DE Human apoptosis regulator (APRG) polypeptide (Incyte ID. 3102521CD1).
 XX KW APRG; apoptosis regulator; cytosolic; antiatherosclerotic; osteopathic;
 KW antiarteriosclerotic; hepatotropic; antiproliferative; antidiabetic; human;
 KW antiallergic; antitumorigenic; antidiabetic; antihypertensive; cancer;
 KW immunosuppressive; dematological; anticancer; nephrotropic; ophthalmological;
 KW antidiabetic; antibacterial; virucide; antiparasitic; prozoacide;
 KW tranquilizer; vulnery; gynecological; vasotropic; gene therapy.
 XX OS Homo sapiens.
 XX PN MO200192527-A2.
 XX PD 06-DEC-2001.
 XX PF 30-MAY-2001; 2001MO-US017581.
 XX PR 01-JUN-2000; 2000US-0209407P.
 XX PR 30-NOV-2000; 2000US-0250326P.
 XX PA (INCY-) INCYTE GENOMICS INC.
 XX PI Tang YT, Azimzai Y, Yue H, Burford N, Ding L, Elliott VS;
 XX PI Patterson C, Baughn MR;
 XX DR WPI; 2002-114350/15.
 XX DR N-PSDB; ABA94362.
 XX PT Novel human apoptosis regulator polypeptides and polynucleotides for
 PT diagnosing, preventing, treating cell proliferative, immunological and
 PT reproductive disorders and for identifying modulators of therapeutic use.
 XX PS Claim 1; Page 97-98; 103pp; English.
 XX CC The invention provides human apoptosis regulator (APRG) polypeptides and


```

QY 361 GAGAGCAACCTGGCTGCTGGGCAACTCTGCGCTGCTGGCCGCGACGACTGCTG 420
DB 80 GluSerAsnLeuArgLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 99
QY 421 CCGGCACTGGCGGCAAGCGCGCGCGCGAGTGTCTCCAGAACGCTATGCTATGGACC 480
DB 100 ProHisLeuAlaArgLysArgArgArgProValSerProGluArgLysSerTYRGLYThr 119
QY 481 TCCAGCTCTCAAGAGAGAGAGGAGTGGTGGCGCTGGCGCGCTGAGTGAAGAGTCT 540
DB 120 SerSerSerSerLysArgGlnGlySerCysArgArgArgArgGlnGlnSerSerSer 139
QY 541 GCAAAATTCAGAGAGGCTGAGGAGAGAGAGAGCTCCCGCCCAACCAAGCGGAGCGG 600
DB 140 AlaAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 159
QY 601 AGTCGGGCGCGCGCGCGCTGCTGCTGCAAGCGCGCGAGAGAGGCGCGCGCGCGACC 660
DB 160 SerArgGlyArgProSerArgLysGlyAlaArgArgArgArgArgArgArgArgArg 179
QY 661 CACGAGAGTCAGAGCGCGCGCGCGCGCTTCTGTAAGCAAGTGAAGTGAAGTGAAGT 720
DB 180 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGln 199
QY 721 CTCGGGCTTCGAGAGAGTACTGCGAGCATGAGCGCGCGCGCTTGGAGAGCGGCTGATCC 780
DB 200 LeuArgValArgAlaGlyArgLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 219
QY 781 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTTGGCGCGCGCGCGCGCGCGCTG 840
DB 220 ArgArgProGlnAlaLeuAlaArgGlnLeuAlaArgValPheGlnAlaThrAlaValLeu 239
QY 841 CGCTCAAGGAGACCTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
DB 240 ArgSerArgAspLeuGlySerValValCysAspLysPheSerGlnLeuSerTYRLeu 259
QY 901 GAGCGCTTGGGCGCGCGCTGAGTGGCGCGCGCTGCTGAGCGCGCGCGCGCTGCTTC 960
DB 260 AspAlaPheTrpGlyArgLysArgLysArgLysArgLysArgLysArgLysArgLys 279
QY 961 CTGACTGAGGCGCTGCGAGAGCGCTGTGGCGCGCGCGCGCTGTTCGCTGCTGCTG 1020
DB 280 LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuValSerVal 299
QY 1021 GATGAGCTACTATGAGCTGTGGCGCGCGCGCGCTGTTCGCTGATGAGAGAGAGGGGG 1080
DB 300 AspGlnAlaAspTYRGLYArgArgArgArgLeuLeuLeuLeuLeuLeuLeuLeu 319
QY 1081 CCGCGCGCGCGAGAGCGCTCC 1101
DB 320 ArgArgProThrGlnAlaSer 326

```

```

RESULT 7
AAE24860
ID AAE24860 standard; protein: 318 AA.
AC AAE24860;
XX
XX 22-OCT-2002 (first entry)
XX
XX Human DED4 (death effector domain) protein.
XX
XX Human; death domain; DD; death effector domain; DED; Chlamydia infection;
XX NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
XX inflammation; allergy; autoimmunity; allograft rejection; cell division;
XX immune-based pathology; fibrosis; arthritis; graft versus host disease;
XX immunosuppressive; gene therapy; antisense therapy.
XX Homo sapiens.
XX Key
XX Location/Qualifiers
XX 53..74
FT Region

```

```

FT /note= "Nuclear localisation sequence"
XX
XX MO200240680-A2.
XX
XX 23-MAY-2002.
XX
XX 15-NOV-2001; 2001MO-US044844.
XX
XX 17-NOV-2000; 2000US-00715893.
XX 29-JUN-2001; 2001US-0301889P.
XX
XX (BURN-) BURNHAM INST.
XX
XX Read JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
XX Stenmer-Liwen F;
XX WPI; 2002-500222/53.
XX N-PSDB; AAD40080.
XX
XX New polypeptide comprising a death domain or death effector domain,
XX useful for discovery of drugs that suppress infection, inflammation,
XX allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
XX Claim 3; Page 186-187; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress responses, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reactive with CTDD DD of C. trachomatis, C.
XX muridarum, C. pneumoniae, and C. piteiaci or a nucleic acid encoding the
XX CTDD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the oncogenic proteins and for
XX treating bacterial infections by modulating the activity of bacterial
XX proteins. The protein and antibody specific for it are useful for
XX discovery of drugs that suppress infection, inflammation, allergy,
XX sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX is useful for treating immune-based pathologies, pathologies associated
XX with cell division, inflammatory diseases such as sepsis, fibrosis,
XX arthritis, graft versus host disease. The invention is used in antisense
XX therapy and gene therapy. The present sequence is human DED4 protein
XX
XX SQ Sequence 318 AA;

```

```

Alignment Scores:
Pred. No.: 3 86e-214 Length: 318
Score: 239.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.55% Indels: 0
DB: 5 Gaps: 0

```

```

US-10-030-271-3 (1-1883) x AAE24860 (1-318)
QY 361 GAGAGCAACCTGGCTGCTGGGCAACTCTGCGCTGCTGGCCGCGACGACTGCTG 420
DB 80 GluSerAsnLeuArgLeuGlyGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 99
QY 421 CCGGCACTGGCGGCAAGCGCGCGCGCGAGTGTCTCCAGAACGCTATGCTATGGACC 480
DB 100 ProHisLeuAlaArgLysArgArgArgProValSerProGluArgLysSerTYRGLYThr 119
QY 481 TCCAGCTCTCAAGAGAGAGAGGAGTGGTGGCGCTGGCGCGCTGAGTGAAGAGTCT 540

```

Db	120	SerSerSerSerLysArgThrGlnGlySerCysArgAlaGArgGlnSerSerSer	139
Qy	541	GCAAAATTCACAGAGGTCAGTGGAGACAGGCTCCCCCAACCAAGCGGACGCGCG	600
Db	140	AlaIasnSerGlnGlnGlyGlnTrpGlnThrGlySerProProThrLysArgGlnArgArg	159
Qy	601	AGTGGGGCCGGCCAGTGTGTGGCCAGACGGCGCGGAGAGGGGCCCAAGCCGACCC	660
Db	160	SerArgGlyArgProSerGlyGlyAlaArgArgArgAlaGArgGlyAlaProAlaLaPro	179
Qy	661	CAGACAGCAGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAAAGTACCTGTGACATCCG	720
Db	180	GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlyLysValThrCysAspIleArg	199
Qy	721	CTCCGGGTTGAGCAGAGACTGTCGAGCATGGCCAGCTTGGAGACAGGGCGTGGCATCC	780
Db	200	LeuArgValArgAlaGlnGlyTyrCysGlnIleHisGlyProAlaLeuGlnGlnGlyValAlaSer	219
Qy	781	CGCGGGCCCCCAGGCGCGCTGGCCGCGACGTGACCGTGTGGGACAGCCACCGCAGTCTG	840
Db	220	ArgArgProGlnAlaLeuAlaArgGlnLeuAspValIleGlyGlnAlaThrAlaValIleu	239
Qy	841	CGCTCAAGAGGACCTGGGCTCTGTGGTTTGTGACATCAAGTTCTCAGAGCTCTCCATCTG	900
Db	240	ArgSerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrIleu	259
Qy	901	GACGCGTTCTGGGGGCGACTTACCTGAGTGGCGCCCTGCTGACAGGCCCTGCGGGCGCTGTC	960
Db	260	AspAlaPheTrpGlyAspTyrIleuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPhe	279
Qy	961	CTGACTGAGGCCCTTCGCGAGAGGCTGTGGCCGCGGAGGCTGTTCGCTGCTGTCAGTGTG	1022
Db	280	LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerVal	299
Qy	1021	GATGAGCGTGACTATGAGCGTGGCCGGCGCGCTGTTGATGATGAGAGAGAGAGG	1077
Db	300	AspGlnAlaAspTyrGlnAlaGlyArgArgArgLeuLeuLeuMetGlnGlnGlnGly	318
RESULT 8			
AAE38903	ID	AAE38903 standard; protein; 318 AA.	
AC	AAE38903;		
XX	XX		
DT	18-DEC-2003 (first entry)		
XX	XX		
DE	Human DED4 protein.		
XX	XX		
KW	Human, death Domain; DB, death effector domain; DED; cell proliferation;		
KW	Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;		
KW	neural growth factor receptor-interacting domain; cell adhesion;		
KW	vasotropic; microbial infection; inflammation; allograft rejection; CTD;		
KW	cell stress response; benign prostatic hypertrophy; antibacterial; NMD;		
KW	apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;		
KW	neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;		
KW	keloid.		
OS	Homo sapiens.		
XX	XX		
FT	Key	Location/Qualifiers	
FT	Domain	12..112	
FT	Peptide	/note="Death effector domain of DED-4"	
FT	FT	53..74	
FT	FT	/note="Nuclear localisation sequence"	
XX	XX		
XX	US2003049702-A1.		
XX	XX		
PD	13-MAR-2003.		
XX	XX		
PE	15-NOV-2001; 2001US-00001254.		
XX	XX		
PR	17-NOV-2000; 2000US-00715893.		

PR 17-NOV-2000; 2000US-0367360P.
 DR 29-JUN-2001; 2001US-0301889P.
 PA (REED/) REED J C.
 PA (GODZ/) GODZIK A.
 PA (PAWL/) PAWLOWSKI K.
 PA (FIOR/) FIORENTINO L.
 PA (LEES/) LEE S H.
 PA (ROTH/) ROTH W.
 PA (STEN/) STENNER-LIEWEN F.
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Steenner-Liewen F;
 DR WPI: 2002-500222/53.
 XX N-PSDB; AAD59062.
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
 PS Claim 2; Page 44-45; 99p; English.
 CC The present invention provides novel death domain (DD) and death effector
 CC domain (DED) proteins and nucleic acids encoding them. The invention also
 CC provides death domain containing protein such as Chlamydia trachomatis
 CC death domain containing protein (CTDD) DD and neutral growth factor
 CC receptor-interacting death domain (NIDD) DD. The invention is useful for
 CC identifying a binding agent (e.g. protein or drug) that binds a DD, DED
 CC or NB-ARC domain from DAP3, IRAK4, CTMD, DED4 or NIDD with a candidate
 CC binding agent and identifying an effective agent (e.g. protein or drug)
 CC that modulates the association of a DD, DED or NB-ARC domain with protein
 CC that binds the DD, DED or NB-ARC domain. The invention is also useful for
 CC modulating the level of cell process such as apoptosis, cell adhesion,
 CC cell proliferation, cell stress responses, responses to microbial
 CC infection and B cell immunoglobulin class switching. DDs, DEDs and NB-ARC
 CC domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are
 CC useful for discovery of drugs that suppress infection, autoimmunity,
 CC inflammation, allergy, allograft rejection, sepsis and other diseases.
 CC DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
 CC autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
 CC hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
 CC inflammatory hyperplasia and smooth muscle cell proliferation in arteries
 CC following balloon angioplasty (restenosis). The invention is also used in
 CC antibody therapy and gene therapy. The present sequence is human DED4
 CC protein. The DED4 gene is located on chromosome 19
 CC XX
 SQ Sequence 318 AA;
 Alignment Scores:
 Pred. No.: 3,866-214 Length: 318
 Score: 239.00 Matches: 239
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 38.55% Indels: 0
 Gaps: 0
 US-10-030-271-3 (1-1883) x AAE38903 (1-318)
 QY 361 GAGAGCAACCTGGCGCTGCTGGGGCAACTCTCGCGCGTCTGGCCCCGACGACCTGCTG 420
 DB 80 GUGSERAAALTAAGLLEUAGLLEUAGLLEUAGLLEUAGLLEUAGLLEUAGLLEUAGLLEU 99
 QY 421 CGCGACCTGGCGCGGCAAGCGGCGCGCGCGAGTGTCTCCAGAGCGCTATAGTCGACCC 480
 DB 100 ProhibisLeuAlaArgLysArgArgArgProValSerProGluArgTyrSerTyrGlyThr 119
 QY 481 TCAGACTTTCAAAGAGACAGAGGGTAGCTGCGCGCGCTGGCGAGTCAAGCACTTCT 540
 DB 120 SerSerSerSerLysArgThrGluGlySerCysAlaGArgArgArgGlnSerSerSerSer 139
 QY 541 GCAAAATTTCAGCAGCGGTCACTGGGAGACAGGCTCTCCCCCAACCAAGCGGCGCGCGG 600

Dh 140 AlaanserGlnGlnGlnTTPGlnThrGlySerProPorthLysArgGlnArg 159
Qy 601 AGTCGGGCGCCAGTGTGTGTCCAGCGGCGGAGAGGGGCCCGCCAGCC 660
Dh 160 SerArgGlyArgProSerGlyGlyAlaArgArgArgArgGlyAlaProAlaPro 179
Qy 661 CAGCAGAGTCAGAGCCCGCCAGACCTTCTCTGAAGGCAAGTGACCTGTGAC 720
Dh 180 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlyValThrCysAspIleArg 199
Qy 721 CTCGGGTTTCAGACAGACTACTGCGAGATATGGCCAGCTTGGAGCGGCGTCATCC 780
Dh 200 LeuArgValArgAlaGlnLysCysGlnLysIleProAlaLeuGlnGlnGlyValAlaSer 219
Qy 781 CGGCGGCGCCAGCGCTGCGCGGAGAGTGAACGTTGTGGCGAGCGCCAGCGAGTGTG 840
Dh 220 ArgArgProGlnAlaLeuAlaArgGlnLeuAspAlaPheGlyGlnAlaThrAlaValLeu 239
Qy 841 CGCTCAAGGAGCTGGGCTGTGGTTTGACATCAAGTTCTGAGAGCTTCTATCTG 900
Dh 240 ArgSerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeuSerTyrLeu 259
Qy 901 GACCCCTTTCGGGCGGACTACTGAGTGGCCCTGCTGCGAGCCCTCGGGCGGTGTC 960
Dh 260 AspAlaPheTyrGlyAspTyrLeuSerGlyAlaLeuLeuGlnAlaLeuArgGlyValPhe 279
Qy 961 CTGACTGAGCGCTGCGAGAGCTGTGGCGGAGGAGCTGTTCGCTGCTGTCAGTGTG 1020
Dh 280 LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerVal 299
Qy 1021 GATGAGGCTGACTATGAGGCTGCGCGCGCGCTGTGTGATGAGAGGAAAGG 1077
Dh 300 AspGlnAlaAspTyrGlnAlaGlyArgArgArgLeuLeuMetGlnGlnGlnGly 318

RESULT 9
AAB58420
ID AAB58420 standard; protein; 319 AA.

XX AAB58420;
AC 14-MAR-2001 (first entry)
XX 14-MAR-2001 (first entry)
XX Lung cancer associated polypeptide sequence SEQ ID 758.

DE Human; lung cancer associated protein; neuroprotective; cytostatic;
XX cardiactive; immunomodulatory; muscular active; vulnery;
KM gastrointestinal; nephrotoxic; antineoplastic; gynecological;
KM antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
KM proliferative disorder; wound healing; infectious disease.

OS Homo sapiens.
XX W020005180-A2.
XX 21-SEP-2000.
XX 08-MAR-2000; 2000WO-US005918.
XX 12-MAR-1999; 99US-0124270P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.

XX Ruben SM;
XX WPI; 2000-587514/55.
XX N-PSDB; AAF18296.
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.

XX Claim 11; Page 1275-1276; 1425bp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardiactive;
CC immunomodulatory; muscular active general; vulnery; gastrointestinal
CC general; nephrotoxic; antineoplastic; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences

Sequence 319 AA;

Alignment Scores:
Pred. No.: 2,16e-181 Length: 319
Score: 204.00 Matches: 304
Percent Similarity: 99.67% Conservative: 0
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 32.90% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x AAB58420 (1-319)

Qy 28 AAGATCCAAACAAAGTGGCTGCGGCGCCAGAGTATCGAGAGCCGAAATTCGAC 87
Dh 9 LysIleGlnAsnLysTyrPheArgProSerProArgSerHisArgThrProGlnSerGly 28
Qy 88 CGGGTTTGAAGCTTGTTCGCTTCCTCCCGGGAATGAGCGCTATCCGGGTGACCCCG 147
Dh 29 ArgValLeuSerLeuPheArgLeuProProProGlyMetAlaLeuSerGlySerThrPro 48
Qy 148 GCCCGTGTGCGGAGAGAGATGAGTGCCTGAGATCAACGGAGATGCTGCTTACCGT 207
Dh 49 AlaProCysTyrGlnGlnLeuArgGlnCysLeuAspTyrTyrGlyMetLeuSerLeuHisArg 68
Qy 208 ATGTTGAGGTGGTGGCGGAGCACTGACCGAGTGGAGTGGAGTCTTGGACCTTCTG 267
Dh 69 MetheGlnValValGlyGlnLeuThrGlnCysLeuGlnLeuLeuAlaPheLeu 88
Qy 268 CTGATGAGGCTCCTGCGCGCGCGAGGCTTACCCCGGCGCCAGCGGCTGAGGCTC 327
Dh 89 LeuAspGlnAlaProGlnAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGlnLeu 108
Qy 328 CTGCTGAGGCTGAGAGCGCCCGGCAATGCGGCGAGAGCACTTGGCGCTTGGGCGA 387
Dh 109 LeuLeuGlnLeuGlnLysArgGlyGlnCysAspIleSerAsnLeuArgLeuGlyGln 128
Qy 388 CTCCTGCGGCTGAGCGCGCGCACGACTGTGCGGAGCTGCGGCGGAGCGGCGCGG 447
Dh 129 LeuLeuArgValLeuAlaArgHisAspLeuLeuProHisLeuAlaArgLysArgArgArg 148
Qy 448 CCAAGTCTCCAGAACGCTATAGCTATGACACTTCAGCTTTCMAAGAGAGAGAGGT 507
Dh 149 ProValSerProGlnArgTyrSerTyrGlyThrSerSerSerSerLysArgThrGlnGly 168
Qy 508 AGCTGCGGTGCGCGGTGCGGAGTCAAGAGTTTTCGCAATTTCTGACGAGGTCACTGGAG 567
Dh 169 SerCysArgArgArgArgGlnSerSerSerSerSerValAsnSerGlnGlnGlnTyrGln 188
Qy 568 ACAGGCTCCCGCCCAACCAAGCGGAGCGGAGGAGTGGGGCGGCGCCAGAGTGGTGGCC 627
Dh 189 ThrGlySerProPorthLysArgGlnArgArgSerArgLysArgProSerGlyGlyAla 208
Qy 628 AGACGGCGGAGAGGGGCCCGACCGGACCCCGAGGAGAGTGCAGAGCCCGCAGACT 687

Db 209 ArgArgArgArgGlyAlaProAlaAlaProGlnGlnGlnSerGluProAlaArgPro 228
QY 688 TCCTCTAGAGCAAGTACCTGTGACATCCGGCTCCGGTTCCAGCAGAGTACTGGCAG 747
Db 229 SerSerGluGlyValThrCysAspIleArgLeuArgValArgAlaGluTyrCysGlu 248
QY 748 CATGGGCGACCTTGGAGCAGGGGCGTGGCATCCCGGGGGCCCAAGGGGCGTGGCGGCGAG 807
Db 249 HisGlyProAlaLeuGlnGlnGlyValAlaSerArgArgProGlnAlaLeuAlaArgGln 268
QY 808 CTGACCTGTTGGGAGAGCCAGCCAGTGGCTGGCTCAAGGAGCTGGGCTGTGGTT 867
Db 269 LeuAspValPheGlyGlnAlaThrAlaValLeuArgSerArgAspLeuGlySerValVal 288
QY 868 TGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGCGGACTACTGAGT 927
Db 289 CysAspIleLysPheSerGluLeuSerTyrLeuAspAlaPheTfpgIAspTyrLeuSer 308
QY 928 GGGCGCTGTGTCGAG 942
Db 309 GlyAlaLeuLeuGln 313

RESULT 10
AAM25705
ID AAM25705 standard; protein; 242 AA.
XX AAM25705;
XX AAM25705;
DT 16-OCT-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:1220.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; vitruide;
KW anti-HIV; fungicide; antitumigen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnerary; antiulcer; osteopathic; eczema;
KW dermatological; antiallergic; antiaethmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
OS Homo sapiens.
XX
PN MO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000MO-US035017.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-0048725.
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457603/49.
DR N-PSDB; AAH99646.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
PS Claim 20; Page 253; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to

CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiallergic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; vitruide; anti-HIV; fungicide; antitumigen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnerary;
CC antitumor; osteopathic; dermatological; antiallergic; antieuthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX
SQ Sequence 242 AA;
XX

Alignment Scores:
Pred. No.: 2 13e-137 length: 242
Score: 157.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.32% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x AAM25705 (1-242)
QY 631 CGGCGGCGAGAGGGGCGCCAGCCGACCCAGCAGAGTCAAGCCCGGCACTTCC 690
Db 86 ArgArgArgArgGlyAlaProAlaAlaProGlnGlnGlnSerGluProAlaArgProSer 105
QY 691 TCGAAGGCAAGTACCTGTGACATCCGCTCCGGTTGGAGCAGTACTGGCAGCAT 750
Db 106 SerGluGlyValThrCysAspIleArgLeuArgValArgAlaGluTyrCysGluHis 125
QY 751 GGGCGACCTTGGAGCAGGGGCGTGGCATCCGGGCGGCCAGGCGCTGGGCGGCACTG 810
Db 126 GlyProAlaLeuGlnGlnGlyValAlaSerArgArgProGlnAlaLeuAlaArgGlnLeu 145
QY 811 GACGTGTTGGGAGGCGCACCGCAGTGTCCGCTCAAGGAGCCTGGCTGTGGTTGT 870
Db 146 AspValPheGlyGlnAlaThrAlaValLeuArgSerArgAspLeuGlySerValValCys 165
QY 871 GACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCTTCTGGGGCGCACTACGATGGC 930
Db 166 AspIleLysPheSerGluLeuSerTyrLeuAspAlaPheTfpgIAspTyrLeuSerGly 185
QY 931 GCCCTGTGAGAGGCGCGGGGCGTGTCTGACTGAGGCGCTGGCAGAGGCTGTGGGC 990
Db 186 AlaLeuLeuGlnAlaLeuArgGlyValAlaPheLeuThrGlnAlaLeuArgGlnAlaValGly 205
QY 991 CGGAGGCTGTGGCTGTGCTGTGCTGATGTGATGAGGCTGACTATGAGGCTGGCGCGC 1050
Db 206 ArgGlnAlaValArgLeuLeuValSerValAspGlnAlaPheTyrGlnAlaGluArgArg 225
QY 1051 CGCTGTTGCTGATGAGAGAGAAAGGGGGCGGCGCCGACAGAGGCTTCC 1101
Db 226 ArgLeuLeuLeuMetGluGlnGlnGlyArgArgProThrGlnAlaSer 242

RESULT 11
ABB06038
ID ABB06038 standard; protein; 361 AA.
XX ABB06038;
XX ABB06038;
DT 10-MAY-2002 (first entry)
XX

DE Human NS protein sequence SEQ ID NO:130.

XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 XX antineumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 XX vasotropic; antiatherosclerotic; antiinflammatory; dermatological;
 XX anorectic; muscular; antinfertility; cardiovascular; anticoagulant;
 XX antifibrinolytic; hypotensive; antiaesthetic; immunomodulator; cardiant;
 XX anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
 XX gastrointestinal; virucide; antitumor; cerebroprotective; nootropic;
 XX contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 XX endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 XX rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 XX inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 XX ischaemia; asthma; immune disease; coagulation disease; hypertension;
 XX diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 XX gastric ulcer; Alzheimer's disease.
 OS Homo sapiens.
 XX MO200206315-A2.
 PN 24-JAN-2002.
 XX 17-JUL-2001; 2001WO-IL000653.
 XX 18-JUL-2000; 2000IL-00137345.
 PR 15-DEC-2000; 2000IL-00140354.
 XX (COMP-) COMPUGEN LTD.
 PA Mintz L, Freilich S, Bernstein J;
 PI WPI: 2002-155037/20.
 DR N-PSDB: ABL39692.
 XX One hundred and twenty eight novel nucleic acid sequences, useful for
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
 XX Claim 6; Page 149-151; 29opp; English.

XX ABL39691 to ABL39818 represent novel human nucleic acid sequences
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
 CC antineumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
 CC vasotropic, antiatherosclerotic, antiinflammatory, dermatological,
 CC anorectic, muscular, anti-HIV, antinfertility, cardiovascular,
 CC anticoagulant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, neuroleptic,
 CC gastrointestinal, virucide, antitumor, cerebroprotective, nootropic;
 CC contraceptive and vaccine activities. The NS can be used in vaccines,
 CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
 CC antibodies from the present invention can be used for treating and
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
 CC Alzheimer's disease and as a contraceptive
 XX
 XX Sequence 361 AA:

Alignment Scores:
 Pred. No.: 1,3e-105 Length: 361
 Score: 123.00 Matches: 149
 Percent Similarity: 98.68% Conservative: 0
 Best Local Similarity: 98.68% Mismatches: 0
 Query Match: 19.84% Indels: 2
 DB: 5 Gaps: 0

US-10-030-271-3 (1-1883) x ABB06038 (1-361)

QY 597 GCGAGTCCGGGCGCCAGTGGTGGCCAGACGGCGCGGAGAGGGCCCGCCAGCGC 656
 Db 186 AAGluserglProAlaGlnTPRCyGlnhrAlaAlaGluArg-AlaProAlaAl 205
 QY 657 ACCCGACGAGTCCAGAGCCCGCCAGACCTTCTCTGAAAGGCAAGTACCTGTGACAT 716
 Db 205 aProGlnGlnGlnSerlunProAlaArgProSerSerGlnGlyValThrCysAepII 225
 QY 717 CCGGCTCCGGGTTTCAGACAGTACCTGCGAGCATGGCCAGCCCTTGAGCAGGGCGTGGC 776
 Db 225 eArgLunArgValArgAlaGluTyrCysGluHISGlyProAlaLunGlnGlnValAl 245
 QY 777 ATCCCGCGCGCCCGCCAGCGCTGGCGCGCGAGCTGACGTGTTGGCAGGCGCCAGT 836
 Db 245 aSerArgArgProGlnAlaLunAlaArgGlnLunAepValPheGlyGlnAlaThrAlaVal 265
 QY 837 GCTGCGCTCAAGGAGCCTGGGCTCTGTGTGTGTGATCATGATTCAGAGCTTCTCTTA 896
 Db 265 lLunArgSerArgAspLunGlnSerValValCysAepIIeLepHeserGlnLunSerTy 285
 QY 897 TCTGGAGCCTTCTGGGGCGGACTACCTGAGTGGCGCCCTGAGGCGCCGCGGGCGT 956
 Db 285 rLunAspAlaPheTrpGlyAspTyLunSerGlyAlaLunLunGlnAlaLunArgGlyVal 305
 QY 957 GTTCTGACTGAGGCGCTGGCGAGAGCTGGCGCGGAGGCTTTCGCTGTGTGACG 1016
 Db 305 lPheLunThrGlnAlaLunArgGlnAlaValGlyArgGlu-LunPheAlaCysTrpSerV 325
 QY 1017 TGTGATGAGGCTGACTATGAGGCTGGCC 1045
 Db 325 aTrpMetArgLunThrMetArgLunAla 334
 RESULT 12
 ID ABB06039 standard; protein: 368 AA.
 AC ABB06039;
 XX 10-MAY-2002 (first entry)
 DT
 XX Human NS protein sequence SEQ ID NO:131.
 DE
 XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;
 KW antineumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
 KW vasotropic; antiatherosclerotic; antiinflammatory; dermatological;
 KW anorectic; muscular; antinfertility; cardiovascular; anticoagulant;
 KW antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiant;
 KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
 KW gastrointestinal; virucide; antitumor; cerebroprotective; nootropic;
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
 KW ischaemia; asthma; immune disease; coagulation disease; hypertension;
 KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
 KW gastric ulcer; Alzheimer's disease.
 OS Homo sapiens.
 XX MO200206315-A2.
 PN 24-JAN-2002.
 XX 17-JUL-2001; 2001WO-IL000653.
 XX 18-JUL-2000; 2000IL-00137345.
 PR 15-DEC-2000; 2000IL-00140354.
 XX (COMP-) COMPUGEN LTD.
 PA Mintz L, Freilich S, Bernstein J;
 PI

Query Match: 14.19% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x AAM41591 (1-217)

QY 37 AACAGAGGCTGGCGCCCTGGCCAGAGATCATGGAACCCAGAAATCTGGCCGGGTTCTG 96
DB 1 ASHLYSTRPHEUATGPROSERPROARGSERISARTRHPRGGUSERGLYATGVALLEU 20
QY 97 AGCTTGTCGCGCTCCCTCCCGGAGATGGCGGTATCCGGGATGGAGCCCGCCCGCTGTC 156
DB 21 SerleupheargleupProProProGlyMethAlaLeuSerGlySerThrProAlaProCys 40
QY 157 TGGGAGAGAGATGAGTGCCTGAGTACTACCGGATGCTGTGCTTCACCGTATGTCAG 216
DB 41 TrpGluGluAhpGluGluCysLeuAhpTrpTrpGlyMetLeuSerLeuHISArgMetPheGlu 60
QY 217 GTGTGGCGCGGCACTGACCGAGATGGAGGTGGAGTCCCTGGGCTTTGCTGGATGAG 276
DB 61 ValValGlyGlyGlnLeuThrGlnCysGlnLeuGlnLeuAlaPheLeuLeuAhpGln 80
QY 277 GCTCCGCGCGCGCGCGGAGGCTTA 300
DB 81 AlaProGlyAlaAlaGlyGlyLeu 88

RESULT 14

AAM39805 ID AAM39805 standard; protein: 277 AA.

AC AAM39805;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2950.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukemia.

OS Homo sapiens.

PN WO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00486725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

XX (HYSB-) HYSBQ INC.

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.
XX N-PSDB; AAI58961.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Example 4; SEQ ID NO 2950; 10078BP; English.
PS

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous system injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilization of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX Sequence 277 AA;

QY Alignment Scores:

Score: 1,486-62 Length: 277
Percent Similarity: 77.00 Matches: 77
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.42% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x AAM39805 (1-277)

QY 1518 ACCAGAACCAACTGGGCTTGAAGCAGATGAGGAGCTGGAGCTGGGCTGACATGGG 1577

DB 190 ThrThrThrAsnLeuGlnPheSerHisIleSerGlyHisTrpSerTrpGlyAlaHisGly 209

QY 1578 GCTGCTCACCCTGCGCCACATCTCCAGCGAGCGGCGCTGCCAGCTTCAATTAC 1637

DB 210 AlaCysSerProCysProHisIleSerSerGlnProGlyProCysProAlaSerIleTrp 229

QY 1638 AGACCTGACTCTCTCCTCACTTCCCTCCCTGCTGTCAGAGCTGAACATGACTTGCACCTTG 1697

DB 230 ArgProAspSerProHisLeuProProCysCysProIleuLeuHisIleAhpLeuHisLeu 249

QY 1698 GATGTCACTGAGATGTACATGAGGAGTGTATGGCAGCATCATACCAAG 1748

DB 250 AspValThrTrpSerValThrTrpGluCysGlySerIleIleProArg 266

RESULT 15

AAE24855 ID AAE24855 standard; protein: 101 AA.

XX AAE24855;

DT 22-OCT-2002 (first entry)

DE Human DED4 DED (death effector domain) protein.

XX Human; death domain; DD; death effector domain; DED; Chlamydia infection;
XX NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
XX inflammation; allergy; autoimmunity; allograft rejection; cell division;
XX immune-based pathology; fibrosis; arthritis; graft versus host disease;
XX immunosuppressive; gene therapy; antisense therapy.

OS Homo sapiens.

PN WO200240680-A2.

XX 23-MAY-2002.

PF 15-NOV-2001; 2001WO-US044844.

PR 17-NOV-2000; 2000US-00715893.

PR 29-JUN-2001; 2001US-0301889P.

XX (BURN-) BURHAM INST.

XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
PI Stenner-Liewen F;
XX WPI; 2002-500222/53.
DR N-PSDB; AAD40075.
XX
PT New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
PS Claim 1; Page 175; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
XX domain (DD), death effector domain (DED) or NB-ARC domain. The invention
XX is useful for identifying a binding agent, preferably a protein or a drug
XX that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
XX domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
XX NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
XX detecting the association of the domain and the candidate binding agent,
XX by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
XX chemical crosslinking, nuclear magnetic resonance (NMR), mass
XX spectroscopy (MS) and FPA. The invention is useful for modulating the
XX level of a cell process such as cell proliferation, cell adhesion, cell
XX stress response, responses to microbial infection and B cell
XX immunoglobulin class switching, in particular apoptosis within a cell.
XX Antibody specifically reactive with CTDD DD of C. trachomatis, C.
XX muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
XX CTDD DD protein is useful for detecting a Chlamydia infection. The
XX invention is useful for modulating the activity of oncogenic proteins,
XX for treating a pathology caused by the oncogenic proteins and for
XX treating bacterial infections by modulating the activity of bacterial
XX proteins. The protein and antibody specific for it are useful for
XX discovery of drugs that suppress infection, inflammation, allergy,
XX sepsis, autoimmunity, allograft rejection and other diseases. The protein
XX is useful for treating immune-based pathologies, pathologies associated
XX with cell division, inflammatory diseases such as sepsis, fibrosis,
XX arthritis, graft versus host disease. The invention is used in antisense
XX therapy and gene therapy. The present sequence is human DED4 DED protein
XX
SQ Sequence 101 AA;
XX
XX Alignment Scores:
XX Pred. No.: 3,91e-53 Length: 101
XX Score: 67.00 Matches: 67
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 10.81% Indels: 0
XX DB: 5 Gaps: 0
XX
XX US-10-030-271-3 (1-1883) x AAE24855 (1-101)
XX
XX QY 157 TGGGAGGAGATGAGTCTGAGCTACTACGCGGATGCTGCTGCTACCGTATGTCGAG 216
XX |
XX DB 1 TGGTGTGAGGAGTCTGAGCTACTACGCGGATGCTGCTGCTACCGTATGTCGAG 216
XX |
XX QY 217 GTGCTGAGGCGGCAACTGACCGAGTGGAGCTGAGCTCTGCGCTTCTGCTGATGAG 276
XX |
XX DB 21 ValValGlyGlyGlnLeuThrGlnCysGlnLeuGlnLeuLeuAlaPheLeuLeuAGln 40
XX |
XX QY 277 GCTTCTGAGGCGGCGGAGGCTTACCGCGCGCGCGCGCTTACGCTGCTGATGAG 336
XX |
XX DB 41 AlaPProIyAlaAlaGlyGlyLeuAlaArgAlaArgserGlyLeuGlnLeuLeuGln 60
XX |
XX QY 337 CTGAGGCGCGCGGCGGAGTGC 357
XX |
XX DB 61 LeuGlnArgArgGlyGlnCys 67
XX
XX RESULT 16
XX ID AAE38898 standard; protein; 101 AA.
XX AC AAE38898;

XX 18-DEC-2003 (first entry)
XX
XX Human DED4 DED protein.
DE
XX
XX Human; death domain; DD; death effector domain; DED; cell proliferation;
KW Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;
KW neural growth factor receptor-interacting death domain; cell adhesion;
KW vascotropic; microbial infection; inflammation; allograft rejection; CTDD;
KW cell stress response; benign prostatic hypertrophy; antibacterial; NIDD;
KW apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;
KW neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;
KW keloid.
XX
XX Homo sapiens.
OS
XX
XX US2003049702-A1.
PN
XX
XX 13-MAR-2003.
PD
XX
XX 15-NOV-2001; 2001US-00001254.
PF
XX
XX 17-NOV-2000; 2000US-00715893.
PR
XX 17-NOV-2000; 2000US-0367360P.
PR
XX 29-JUN-2001; 2001US-0301889P.
PR
XX
XX (REED/) REED J C.
PA (GODZ/) GODZIK A.
PA (PAWL/) PAWLOWSKI K.
PA (FIOR/) FIORENTINO L.
PA (LEES/) LEE S H.
PA (ROTH/) ROTH W.
PA (STEN/) STENNER-LIEWEN F.
XX
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
PI Stenner-Liewen F;
XX WPI; 2002-500222/53.
DR N-PSDB; AAD59057.
XX
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
XX Claim 1; Page 35; 99pp; English.
XX
XX The present invention provides novel death domain (DD) and death effector
XX domain (DED) proteins and nucleic acids encoding them. The invention also
XX provides death domain containing protein such as Chlamydia trachomatis
XX death domain containing protein (CTDD) DD and neural growth factor
XX receptor-interacting death domain (NIDD) DD. The invention is useful for
XX identifying a binding agent (e.g. protein or drug) that binds a DD, DED
XX or NB-ARC domain from DAP3, IRAK4, CTDD, DED4 or NIDD with a candidate
XX binding agent and identifying an effective agent (e.g. protein or drug)
XX that modulates the association of a DD, DED or NB-ARC domain with protein
XX that binds the DD, DED or NB-ARC domain. The invention is also useful for
XX modulating the level of cell process such as apoptosis, cell adhesion,
XX cell proliferation, cell stress responses, responses to microbial
XX infection and B cell immunoglobulin class switching. DEDs, DEDs and NB-ARC
XX domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are
XX useful for discovery of drugs that suppress infection, autoimmunity,
XX inflammation, allergy, allograft rejection, sepsis and other diseases.
XX DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
XX autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
XX hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
XX inflammatory hyperplasia and smooth muscle cell proliferation in arteries
XX following balloon angioplasty (restenosis). The invention is also used in
XX antibody therapy and gene therapy. The present sequence is human DED4 DED
XX protein. The DED4 gene is located on chromosome 1
XX
SQ Sequence 101 AA;
XX
XX Alignment Scores:
XX

Pred. No.: 3,91e-53 Length: 101
 Score: 67.00 Matches: 67
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 10.81% Indels: 0
 DB: 5 Gaps: 0

US-10-030-271-3 (1-1883) x AAE38898 (1-101)

QY 157 TGGAGAGAGATGATGCTGCTGACTACAGGAGATGCTGCTTACCCGATATGTCGAG 216
 DB 1 TTPGLUGLUSPGLUCYSLEUASPITYRTYGLYMETLEUSERLEUHLAARGMETPHEGLU 20
 QY 217 GTGCTGGGCGGCGCACTGACCGAGTGGAGCTGAGCTCTGAGCTTCTGCTGAGAG 276
 DB 21 ValValGlyGlyGlnLeuThrGluCysGluLeuGluLeuLeuAlaPheLeuLeuGlu 40
 QY 277 GCTCTGTCGCGCGCCGAGGCTTAGCCCGGCGCGAGCGGCTTAGAGCTCTGCTGAG 336
 DB 41 AlaProGlyAlaAlaGlyGlyLeuAlaArgAlaArgSerGlyLeuGluLeuLeuGlu 60
 QY 337 CTGGAGCGCGCGGCGGAGTGC 357
 DB 61 LeuGluArgArgGlyGlnCys 67

RESULT 17

ABP59171 ID ABP59171 standard; protein, 105 AA.

AC ABP59171;

DT 25-APR-2003 (first entry)

DE ATP synthetase 11.55.

KM ATP synthetase 11.55; enzyme; cancer; cytostatic; HIV infection;
 anti-HIV.

OS Unidentified.

XX CNI363656-A.

PD 14-AUG-2002.

PF 05-JAN-2001; 2001CN-00105031.

PR 05-JAN-2001; 2001CN-00105031.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

PI Mao Y, Xie Y;

DR WPI; 2002-751770/82.

DR N-PSDB; ABZ70266.

PT Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.

PS Claim 1; Page 26 (disclosure); 32pp; Chinese.

CC The present sequence is the protein sequence for ATP synthetase 11.55.

CC The protein can be used for treating diseases such as cancer and HIV

XX infection

XX Sequence 105 AA;

XX Alignment Scores:

Pred. No.: 1.02e-44 Length: 105
 Score: 58.00 Matches: 58
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 9.35% Indels: 0
 DB: 5 Gaps: 0

US-10-030-271-3 (1-1883) x ABP59171 (1-105)

QY 1201 ATGGGGCTCTGAGGCGCTAAACCATTTCCAGCTGATTTCTTCCAGACTCCTCTACC 1260

DB 1 MetTrpAlaLeuArgProLysProPheProAlaGluPheProSerGlnTrpProThr 20

QY 1261 CCCAGGTGTGCCCCCTTAGCCTCCGAGGCGGGGCTGGGCTGTATCTCAGAAAGGAG 1320

DB 21 ProArgCysAlaProLeuAlaSerGlyGlyGlyTrpAlaCysIleSerGluGlyArg 40

QY 1321 GGCACAGCTACACACTACCAAGAGCCCGCTGCACATTTGATCTTGATTTG 1374

DB 41 GlyThrAlaThrHisSerProLysAlaProLeuHisIleValSerLeuIleLeu 58

RESULT 18

AAB15560 ID AAB15560 standard; peptide; 40 AA.

AC AAB15560;

DT 28-FEB-2001 (first entry)

DE Fragment of apoptosis related protein encoded by gene 1 clone HLDOK36.

KM Cytostatic; antiinflammatory; immunosuppressive; antisclerotic; cardiant;
 virucidal; anti-AIDS; vasectropic; anti-ischaemic; antiparkinsonian;

KM anti-Alzheimer; gene therapy; human; apoptosis; fusion protein; cancer;
 colon; breast; prostate; melanoma; lymphoma; inflammation; herpes;

KM autoimmune disorder; multiple sclerosis; viral infection.

XX Homo sapiens.

PN WO200056752-A2.

XX 28-SEP-2000.

PF 15-MAR-2000; 2000WO-US006642.

PR 24-MAR-1999; 99US-0126018P.

PR 17-JUN-1999; 99US-0139638P.

PR 18-AUG-1999; 99US-0149449P.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Ni J, Young PA;

DR WPI; 2000-587660/55.

PT Nucleic acids encoding human apoptosis associated protein, useful for the
 prevention, treatment and diagnosis of e.g. Alzheimer's and Parkinson's
 disease, inflammation and ischemic injury.

PS Disclosure; Page 17; 273pp; English.

XX The invention relates to the isolation of genes encoding 9 human

CC apoptosis-related proteins. The nucleotide sequences ABA95790-A95798

CC encode the human apoptosis related proteins AAB15551-B15559. This

CC sequence represents a fragment of the protein encoded by the gene given

CC in the descriptor line. The genes can be used to generate fusion proteins

CC by linking to the gene for the human immunoglobulin G Fc (IgG Fc) portion

CC (AAB95799) for increasing the stability of the fusion protein as compared

CC to the human protein only. The gene and encoded protein may be used in

CC the prevention, treatment and diagnosis of diseases associated with in

CC inappropriate apoptosis associated protein expression, e.g. cancer (e.g.

CC colon, breast and prostate cancer, melanomas and lymphomas),

CC inflammation, autoimmune disorders (e.g. multiple sclerosis) and viral

XX infections (e.g. herpes))

XX Sequence 40 AA;

XX Alignment Scores:

Pred. No.: 2.5e-20 Length: 40
 Score: 32.00 Matches: 32

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 5.16%
 DB: 3
 Gaps: 0

US-10-030-271-3 (1-1883) x AAB15560 (1-40)

QY 28 AAGATCCAAAGAGTGGTGGCGCCGCGAGAGTGCAGACGCCAATCTGAC 87
 Db 9 Lye11EGlnahsnlystrleuArgProserProArgserHisArgThrProGlnserGly 28
 QY 88 CGGGTCTGAGCTTGTTCCTCCCTCCCGCCGGA 123
 Db 29 ArgValIeuSerIeuPheArgIeuProProGly 40

RESULT 19

ABP59172
 ID ABP59172 standard; peptide; 15 AA.

AC ABP59172;

DT 25-APR-2003 (first entry)

DE ATP synthetase 11.55 peptide fragment.

KW ATP synthetase 11.55; enzyme; cancer; cytosolic; HIV infection;
 anti-HIV.

OS Unidentified.

PN CN1363656-A.

PD 14-AUG-2002.

PF 05-JAN-2001; 2001CN-00105031.

PR 05-JAN-2001; 2001CN-00105031.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

PI Mao Y, Xie Y;

DR WPI; 2002-751770/82.

PT Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.

PS Example 5; Page 18 (Disclosure); 32pp; Chinese.

CC The present invention relates to ATP synthetase 11.55 (see ABP59171). The

CC protein can be used for treating diseases such as cancer and HIV

CC infection. The present sequence is an N-terminal peptide fragment of the

CC protein, which was used in an example from the invention

XX

XX

XX

XX

XX

XX

XX

XX

XX

AC AAM90108;
 XX 12-APR-1999 (first entry)
 DT
 XX
 DE Human FLAME-2 protein.

KW FLAME-2; FADD-like apoptotic/anti-apoptotic molecule; human; apoptosis;
 KW HIV; infection; Alzheimer's disease; cancer; therapy.

OS Homo sapiens.

FT Key Location/Qualifiers

FT Domain 23..101

FT Domain /note="FADD-DED homology domain"

FT Domain 102..318

FT Domain /note="C-terminal domain"

PN WO9852963-A1.

PD 26-NOV-1998.

PF 20-MAY-1998; 98WO-US010200.

PR 20-MAY-1997; 97US-00859167.

PA (UYJE-) UNIV JEFFERSON THOMAS.

PI Alnemri ES;

DR WPI; 1999-045296/04.

DR N-PSDB; AAV74137.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

SEQ Sequence 318 AA;

Alignment Scores:

Pred. No.: 0.0111 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x AAM90108 (1-318)

QY 706 ACCTGTGACATCCGGGCTTCGAGAGAGTACTGC 744
 Db 192 ThCysAspIleArgIeuArgValArgAlaGluTrpCys 204

RESULT 21

AAM90109

XX

AC AAM90109;
 XX
 DT 12-APR-1999 (first entry)
 XX
 DE Mouse FLAME-2 protein.
 XX
 KM FLAME-2; FADD-like apoptotic/anti-apoptotic molecule; mouse; apoptosis;
 KW HIV; infection; Alzheimer's disease; cancer; therapy.
 XX
 OS Mus sp.
 XX
 PN M09852963-A1.
 XX
 PD 26-NOV-1998.
 XX
 PF 20-MAY-1998; 98MO-US010200.
 XX
 PR 20-MAY-1997; 97US-00859167.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 PI Alnemri ES;
 XX
 DR WPI: 1999-045296/04.
 XX
 DR N-PSDB; AAV74138.
 XX
 PT New isolated FADD-like anti-apoptotic molecules - used to develop
 PT apoptotic and anti-apoptotic agents for treating, e.g. HIV infection,
 PT Alzheimer's disease or neoplastic conditions.
 XX
 PS Example; Page 44-45; 66pp; English.
 XX
 CC This is the amino acid sequence of mouse FLAME-2, or FADD-like
 CC apoptotic/anti-apoptotic molecule 2. FLAME-2 is a novel anti-apoptotic
 CC protein that is similar in structure to FADD, but its C-terminal region
 CC does not have death domain homology. Its structure suggests that it could
 CC be an adaptor molecule for an as yet unidentified signalling complex.
 CC FLAME-2 interacts weakly with Mch4 and Mch5 but does not interact with
 CC FADD. It can abrogate UV-induced apoptosis and to a lesser degree inhibit
 CC Fas/TNFR-induced apoptosis in the same cell line. The FLAME-2 amino acid
 CC sequence was deduced from the nucleotide sequence of an isolated EST (see
 CC AAV74138). Human FLAME-1 and FLAME-2 proteins (see AAM90107-08) are
 CC claimed. Host cells, recombinant vectors, and methods of using FLAME to
 CC identify substrates, activators or inhibitors of FLAME are provided.
 CC FLAME-1, FLAME-2 and agonists can be used to inhibit apoptosis, e.g. for
 CC treating HIV infection or Alzheimer's disease. Inhibitors of the
 CC polypeptides can be used as apoptotic agents
 XX
 SQ Sequence 318 AA;
 XX
 Alignment Scores:
 Pred. No.: 0.0111 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: Gaps: 0
 XX
 US-10-030-271-3 (1-1883) x AAM90109 (1-318)
 QY 706 ACCTGTGACATCCGGCTCCGGGTTTCAGACAGACTACTGC 744
 ID AAV51023
 XX
 AC AAV51023 standard; protein; 318 AA.
 XX
 DT 17-MAR-2000 (first entry)
 XX
 DE Murine DEDD protein.
 XX

XX
 KM DEDD; murine; apoptosis; protein biosynthesis inhibitor; CD95.
 XX
 OS Mus sp.
 XX
 PN DE19825621-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 08-JUN-1998; 98DE-01025621.
 XX
 PR 08-JUN-1998; 98DE-01025621.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Peter M, Krammer P;
 XX
 DR WPI: 2000-063547/06.
 XX
 DR N-PSDB; AAZ43924.
 XX
 PT New protein for regulating apoptosis, particularly for diagnosis or
 PT treatment of e.g. tumors and acquired immune deficiency syndrome.
 XX
 PS Claim 7, Fig 1B; 12pp; German.
 XX
 CC This invention describes a novel DEDD protein (I) for regulating
 CC apoptosis which has antitumor and antiviral activity. (I) is expressed
 CC ubiquitously and after induction of the CD95 apoptotic signalling pathway
 CC migrates to the nucleus/nucleolus where it inhibits transcription of
 CC ribosomal DNA, and thus biosynthesis of proteins, including those with
 CC anti-apoptotic activity. (I), or the DNA (II), encoding it, is used for
 CC regulation and diagnostic detection of apoptosis, particularly in cases
 CC of disease, e.g. of the immune system (e.g. acquired immune deficiency
 CC syndrome) and tumors, also for studying apoptosis or its regulation.
 CC Antibodies (Ab) specific for (I) are used to detect/quantify (I),
 CC particularly as a function of time, also for inhibition of (I). (I) is
 CC used to raise Ab and for detection of specific autoantibodies. (II), or
 CC derived primers, can be used to detect expression and organization of the
 CC corresponding gene, also for inducing expression of (I) in vivo or in
 CC selected tissues. This sequence represents the murine DEDD protein
 CC described in the method of the invention
 XX
 SQ Sequence 318 AA;
 XX
 Alignment Scores:
 Pred. No.: 0.0111 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: Gaps: 0
 XX
 US-10-030-271-3 (1-1883) x AAV51023 (1-318)
 QY 706 ACCTGTGACATCCGGCTCCGGGTTTCAGACAGACTACTGC 744
 ID AAV51022
 XX
 AC AAV51022;
 XX
 DT 17-MAR-2000 (first entry)
 XX
 DE Human DEDD protein.
 XX
 KM DEDD; human; apoptosis; protein biosynthesis inhibitor; CD95.
 XX
 OS Homo sapiens.
 XX
 PN DE19825621-A1.
 XX

XX 09-DEC-1999.
XX 08-JUN-1998; 98DE-01025621.
XX 08-JUN-1998; 98DE-01025621.
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX Peter M, Krammer P;
XX WPI: 2000-063547/06.
XX N-PDB; AA243923.
XX New protein for regulating apoptosis, particularly for diagnosis or
XX treatment of e.g. tumors and acquired immune deficiency syndrome.
XX Claim 1; Fig 1A, 12pp; German.
XX This invention describes a novel DEDD protein (I) for regulating
XX apoptosis which has antitumor and antiviral activity. (I) is expressed
XX ubiquitously and after induction of the CD95 apoptotic signalling pathway
XX migrates to the nucleus/nucleolus where it inhibits transcription of
XX ribosomal DNA, and thus biosynthesis of proteins, including those with
XX anti-apoptotic activity. (I), or the DNA (II), encoding it, is used for
XX regulation and diagnostic detection of apoptosis, particularly in cases
XX of disease, e.g. of the immune system (e.g. acquired immune deficiency
XX syndrome) and tumors, also for studying apoptosis or its regulation.
XX Antibodies (Ab) specific for (I) are used to detect/quantify (I),
XX particularly as a function of time, also for inhibition of (I). (I) is
XX used to raise Ab and for detection of specific autoantibodies. (II), or
XX derived primers, can be used to detect expression and organization of the
XX corresponding gene, also for inducing expression of (I) in vivo or in
XX selected tissues. This sequence represents the human DEDD protein
XX described in the method of the invention
SQ Sequence 318 AA;
Alignment Scores:
Pred. No.: 0.011 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: Gaps: 0
US-10-030-271-3 (1-1883) x AAY51022 (1-318)
QY 706 ACCTGTGACATCCGGCTCCGGGTTCCGACGAGACTACTGC 744
Db 192 ThCysAspIleArgLeuArgValArgAlaGluIuTcYs 204
RESULT 24
AAB93016
ID AAB93016 standard; protein; 318 AA.
AC AAB93016;
XX 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:11773.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX EPI074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX

PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-0018776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX Claim 8; SEQ ID NO 11773; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesising polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
SQ Sequence 318 AA;
Alignment Scores:
Pred. No.: 0.011 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: Gaps: 0
US-10-030-271-3 (1-1883) x AAB93016 (1-318)
QY 706 ACCTGTGACATCCGGCTCCGGGTTCCGACGAGACTACTGC 744
Db 192 ThCysAspIleArgLeuArgValArgAlaGluIuTcYs 204
RESULT 25
AAB94040
ID AAB94040 standard; protein; 318 AA.
AC AAB94040;
XX 26-JUN-2001 (first entry)
XX Human protein sequence SEQ ID NO:14195.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX Homo sapiens.
XX EPI074617-A2.
XX

XX 07-FEB-2001.
 PD
 XX 28-JUL-2000; 2000EP-00116126.
 PF
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Iehli S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX WPI; 2001-318749/34.
 DR
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 PS
 XX Claim 8; SEQ ID NO 14195; 2537bp + Sequence Listing, English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the 5602
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 XX
 SQ Sequence 318 AA;
 XX
 Alignment Scores:
 Pred. No.: 0.0111 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: 4 Gaps: 0
 US-10-030-271-3 (1-1883) x AAB94040 (1-318)
 QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGACAGACTACTGC 744
 Db |||||
 192 ThrcYshspilearlglenAargValArgAlagIdutrycys 204
 RESULT 26
 AAE26087
 ID AAE26087 standard; protein; 318 AA.
 XX
 AC AAE26087;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Human FLAME-2 protein.

XX Human; FADD-like apoptotic/anti-apoptotic protein; Alzheimer's disease;
 KW gene therapy; human immunodeficiency virus; HIV infection; apoptosis;
 KM FLAME-2.
 XX
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Region 1..22
 FT Region /note= "1B-2 region"
 FT Region 23..101
 FT Domain /note= "FADD-DED-Homology region; N-terminal region"
 FT /note= ".318
 FT /note= "C-terminal domain"
 XX
 PN US2002086983-A1.
 XX
 PD 04-JUL-2002.
 XX
 XX 22-AUG-2001; 2001US-00935223.
 PP
 XX 28-OCT-1997; 97US-00959167.
 PR 26-MAR-1999; 99US-00276993.
 PR 28-NOV-2000; 2000US-00723450.
 XX
 PA (UYJE-) UNIV JEFFERSON THOMAS.
 PI Alnemri ES;
 XX WPI; 2002-642259/69.
 DR N-PSDB; AAD43203.
 DR
 XX Novel FADD-like apoptotic/anti-apoptotic proteins useful for inhibiting
 PT apoptosis, treating diseases characterized by apoptosis e.g. HIV
 PT infection and Alzheimer's disease, and for identifying modulators of the
 PT protein.
 XX
 PS Claim 1; Page 16-17; 35pp; English.
 XX
 CC The invention relates to FADD-like apoptotic/anti-apoptotic proteins
 CC (FLAME 1 or 2) and nucleic acid molecules encoding such proteins. FLAME
 CC sequences are useful for inhibiting apoptosis and for gene therapy of
 CC diseases characterised by apoptosis including HIV infection and
 CC Alzheimer's disease. FLAME inhibitors are useful as apoptotic agents and
 CC activators are useful as anti-apoptotic agents. FLAME-1 is useful as a
 CC substrate for caspase in assays to identify caspase inhibitors. The
 CC present sequence is human FLAME-2 protein
 CC
 XX
 SQ Sequence 318 AA;
 XX
 Alignment Scores:
 Pred. No.: 0.0111 Length: 318
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.10% Indels: 0
 DB: 5 Gaps: 0
 US-10-030-271-3 (1-1883) x AAE26087 (1-318)
 QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGACAGACTACTGC 744
 Db |||||
 192 ThrcYshspilearlglenAargValArgAlagIdutrycys 204
 RESULT 27
 AAE26088
 ID AAE26088 standard; protein; 318 AA.
 XX
 AC AAE26088;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Mouse FLAME-2 protein.

XX Mouse; FADD-like apoptotic/anti-apoptotic protein; Alzheimer's disease;
KM gene therapy; human immunodeficiency virus; HIV infection; apoptosis;
KM FLAME-2.
XX
OS Mus sp.
PN US2002086983-A1.
PD 04-JUL-2002.
XX
PF 22-AUG-2001; 2001US-00935223.
XX
PR 28-OCT-1997; 97US-00959167.
PR 26-MAR-1999; 99US-00276993.
XX 28-NOV-2000; 2000US-00723450.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
PI Alnemri ES;
XX
XX WPI; 2002-642259/69.
DR N-PSDB; AAD43204.
XX
XX
PT Novel FADD-like apoptotic/anti-apoptotic proteins useful for inhibiting
PT apoptosis, treating diseases characterized by apoptosis e.g. HIV
PT infection and Alzheimer's disease, and for identifying modulators of the
PT protein.
XX
PS Disclosure; Page 19; 35pp; English.
XX
XX The invention relates to FADD-like apoptotic/anti-apoptotic proteins
CC (FLAME 1 or 2) and nucleic acid molecules encoding such proteins. FLAME
CC sequences are useful for inhibiting apoptosis and for gene therapy of
CC diseases characterised by apoptosis including HIV infection and
CC Alzheimer's disease. FLAME inhibitors are useful as apoptotic agents and
CC activators are useful as anti-apoptotic agents. FLAME-1 is useful as a
CC substrate for caspase in assays to identify caspase inhibitors. The
CC present sequence is mouse FLAME-2 protein
XX
SQ Sequence 318 AA;
XX
XX
XX Alignment Scores:
Pred. No.: 0.0111 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 5 Gaps: 0
XX
US-10-030-271-3 (1-1883) x AAE26088 (1-318)
OY 706 ACCGTGACATCCGGCTCCGGGTTGAGAGAGACTGC 744
DB 192 ThCTyehpilleagLeuArgValArgAlaGluIrrCys 204
XX
RESULT 28
AAM39805
ID AAM39805 standard; protein; 277 AA.
XX
AC AAM39805;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2950.
XX
XX Human; noctropic; immunosuppressant; cyostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX

OS Homo sapiens.
XX
PN WO200153312-A1.
XX
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI58961.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX
PS Example 4; SEQ ID NO 2950; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AA42213) with noctropic,
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
SQ Sequence 277 AA;
XX
XX
XX Alignment Scores:
Pred. No.: 0.0975 Length: 277
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.95% Indels: 0
DB: 4 Gaps: 0
XX
US-10-030-271-3 (1-1883) x AAM39805 (1-277)
OY 1796 CCAAGAGGCTGCTGTTTACTGTTTGGCCCGCANG 1761
DB 265 ProArgArgTrrpLeuTherGlyPheGlyPromet 276
XX
RESULT 29
AAU41693
ID AAU41693 standard; protein; 104 AA.
XX
XX AAU41693;
AC
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #2569.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KW vevitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acne.
 KW
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maionneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59515.
 XX
 PT Propionibacterium acne polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 2888; 1069pp; English.
 XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acne immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), vevitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 104 AA;
 XX
 Alignment Scores:
 Pred. No.: 8.36 Length: 104
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.61% Indels: 0
 DB: 4 Gaps: 0
 US-10-030-271-3 (1-1883) x AAU41693 (1-104)
 QY 134 CCGGCTGACCCCGGCGCCCGTGTGGAGG 163
 Db 60 ProGlyArgProArgProArgAlaGlyArg 69
 RESULT 30
 ABM38212
 ID ABM38212 standard; protein; 104 AA.
 XX

AC ABM38212;
 XX
 DT 20-OCT-2003 (first entry)
 XX
 DE Propionibacterium acne predicted ORF-encoded polypeptide #2888.
 XX
 KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acne.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vailleve-Douglas J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF6444.
 XX
 PT New Propionibacterium acne polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 2888; 1481pp; English.
 XX
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acne protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 104 AA;
 XX
 Alignment Scores:
 Pred. No.: 8.36 Length: 104
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.61% Indels: 0
 DB: 6 Gaps: 0
 US-10-030-271-3 (1-1883) x ABM38212 (1-104)

QY 134 CCGGGTCGACCCCGCGCTGCTGGAGG 163
 |||||
 Db 60 ProGlyArgProArgProArgAlaGlyArg 69
 RESULT 31
 AAU30377
 ID AAU30377 standard; protein; 314 AA.
 AC AAU30377;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Novel human secreted protein #868.
 XX
 KW Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEO INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX
 PS Claim 20; Page 285; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU30377-10-030-271-3 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 314 AA;
 Alignment Scores:
 Pred. No.: 7.12 Length: 314
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.61% Indels: 0
 DB: 4 Gaps: 0
 US-10-030-271-3 (1-1883) x AAU30377 (1-314)
 QY 780 CCGGGCGCCAGCGCTGGCGCGGAGCT 809
 |||||

Db 48 ProAlaAlaProGlyAlaGlyAlaAla 57
 RESULT 32
 AAU55376
 ID AAU55376 standard; protein; 51 AA.
 XX
 AC AAU55376;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #16272.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhactia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 XX
 DR N-PADB; AAS59569.
 XX
 PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 16571; 1069pp; English.
 XX
 CC Sequences AAU3105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 51 AA;
 Alignment Scores:
 Pred. No.: 80 Length: 51
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x AAU55376 (1-51)

Qy 350 CCGGCGGCTCCAGCTCCAGAGAGC 324
 |||||
 DB 25 ProHrgrSerrSerrSerrSerr 33

RESULT 33

AAU41040
 ID AAU41040 standard; protein; 51 AA.

XX AAU41040;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #1936.

XX SAPRO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L.maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AASS9513.

XX Example 1; SEQ ID NO 2235; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX polypeptides. The proteins and their associated DNA sequences are used in

XX the treatment, prevention and diagnosis of medical conditions caused by

XX P. acnes. The disorders include SAPRO syndrome (synovitis, acne,

XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.

XX P. acnes is also involved in infections of bone, joints and the central

XX nervous system, however it is particularly involved in the inflammatory

XX lesions associated with acne vulgaris. A method for detecting the

XX presence or absence of P. acnes in a patient comprises contacting a

XX sample with a binding agent that binds to the proteins of the invention

XX and determining the amount of bound protein in the sample. The

XX specific for P. acnes proteins. These antibodies can be used to

XX downregulate expression and activity of P. acnes polypeptides and

XX therefore treat P. acnes infections. The antibodies may also be used as

XX diagnostic agents for determining P. acnes presence, for example, by

XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for

XX this patent did not form part of the printed specification, but was

XX obtained in electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 51 AA;

XX Alignment Scores:

XX Pred. No.: 80 Length: 51

Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x AAU41040 (1-51)

Qy 269 AGCAGAAAGCCAGAGCTCCAGCTCG 243

DB 40 SerHrgrSerrSerrSerrSerr 48

RESULT 34

ABM37559
 ID ABM37559 standard; protein; 51 AA.

XX ABM37559;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #2235.

XX Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;

XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX Barth B, Vallieve-Douglas J;

XX WPI; 2003-381789/36.

XX N-PSDB; ACF64442.

XX Example 1; SEQ ID NO 2235; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)

XX encoding a Propionibacterium acnes protein. The invention also relates to

XX polynucleotides encoded by the polynucleotides (ABM35624-ABM64536) and to

XX immunogenic fragments of P. acnes polypeptides. The invention

XX additionally encompasses expression vectors and host cells comprising a

XX polynucleotide of the invention; antibodies against polypeptides of the

XX invention; fusion proteins comprising a polypeptide of the invention; a

XX method for stimulating an immune response specific for a P. acnes

XX polypeptide and an isolated T cell population comprising T cells prepared

XX via this method; a vaccine composition (comprising P. acnes polypeptides,

XX polynucleotides, antibodies, fusion proteins, T cell populations, or

XX antigen-presenting cells that express the polypeptide); a method and kit

XX for detecting or determining the presence or absence of P. acnes in a

XX patient; and a method for inhibiting the development of P. acnes in a

XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion

XX proteins, T cell populations or antigen-presenting cells that express the

XX polypeptides are useful for diagnosing, preventing or treating acne

XX vulgaris, or for stimulating an immune response specific for a P. acnes

XX protein. The polynucleotides can also be used as probes or primers for

XX nucleic acid hybridisation. The vaccine composition is useful for the

XX stimulation of an immune response against P. acnes, or for treating acne,

XX and the kit is useful for performing a diagnostic assay. The present

XX sequence represents a polypeptide predicted to be encoded by an ORF (open

CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SO Sequence 51 AA;

Alignment Scores:

Pred. No.:	80	Length:	51
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	6	Gaps:	0

US-10-030-271-3 (1-1883) x ABM51895 (1-51)

QY 269 AGCAGAAAGCCAGAGCTCCAGCTCG 243
Db 40 SerArgLysAlaArgSerSerSerSer 48

RESULT 35
ABM51895
ID ABM51895 standard; protein; 51 AA.
XX
AC ABM51895;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #16571.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vailleve-Douglas J;
XX
XX WPI; 2003-381789/36.
XX N-PDB; ACF64498.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or stimulating an immune response specific for a P. acnes protein.
XX
XX Example 1; SEQ ID NO 16571; 1481bp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM51895-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a

CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC protein, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridization. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SO Sequence 51 AA;

Alignment Scores:

Pred. No.:	80	Length:	51
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	6	Gaps:	0

US-10-030-271-3 (1-1883) x ABM51895 (1-51)

QY 350 CCGCGGCGCTCCAGCTCCAGAGAGC 324
Db 25 ProArgArgSerSerSerSerSerSer 33

RESULT 36
ADN01127
ID ADN01127 standard; protein; 68 AA.
XX
AC ADN01127;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human cell growth, differentiation, and death-associated protein #11.
XX
KW human; cell growth; cell differentiation; cell death; CGDP;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis; bursitis;
KW cirrhosis; hepatitis; polycythaemia vera; psoriasis; cancer;
KW developmental disorder; Cushing's syndrome; hypothyroidism;
KW neurological disorder; epilepsy; stroke; Alzheimer's disease;
KW Pick's disease; Huntington's disease; Parkinson's disease;
KW multiple sclerosis; autoimmune disorder; inflammatory disorder; AIDS;
KW allergy; anaemia; asthma; contact dermatitis; diabetes mellitus;
KW reproductive disorder; infertility; endometriosis; uterine fibroid.
XX
XX Homo sapiens.
XX
XX WO2004031364-A2.
XX
XX 15-APR-2004.
XX
XX 03-OCT-2003; 2003WO-US031441.
XX
XX 03-OCT-2002; 2002US-0416205P.
XX 25-OCT-2002; 2002US-0421521P.
XX 21-NOV-2002; 2002US-0428376P.
XX 23-DEC-2002; 2002US-036258P.
XX 10-JAN-2003; 2003US-0439292P.
XX 13-FEB-2003; 2003US-0447578P.
XX
XX (INCY-) INCYTE CORP.
XX (BURL) BURRILL J D.
XX
XX Elliott VS, Swarnakar A, Tang YT, Yue H, Lu DM, Khare R;
XX Chawla NK, Richardson TW, Margulis JP, Lal PG, Nguyen DB, Lee SY;
XX Tran UK, Bhatia UG, Lee S, Blake JJ, Ho A, Zheng W, Gao J, Tran B;
XX Yang YG, Gietzen KJ, Hafalia AA;

DR WP1: 2004-330172/30.
DR N-PSDB; ADN01152.
XX
XX
XX New isolated polypeptides associated with cell growth, differentiation
PT and death, useful for diagnosing, treating or preventing e.g.
PT atherosclerosis, psoriasis, cancer, Alzheimer's disease, AIDS, anemia,
PT diabetes mellitus or infertility.
XX
XX
PS Claim 1; SEQ ID NO 11; 213bp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC proteins that are associated with cell growth, differentiation, and death
CC (CGSD). The DNA and protein sequences of the invention are useful for
CC diagnosis, treating or preventing disorders associated with aberrant
CC expression of CGSD, such as: cell proliferative disorders (e.g.
CC atherosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis,
CC polycythaemia vera, psoriasis and cancer), developmental disorders (e.g.
CC Cushing's syndrome and hypothyroidism), neurological disorders (e.g.
CC epilepsy, stroke, Alzheimer's disease, Pick's disease, Huntington's
CC disease, Parkinson's disease and multiple sclerosis),
CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, anaemia, asthma,
CC contact dermatitis and diabetes mellitus), and reproductive disorders
CC (e.g. infertility, endometriosis and uterine fibroid). The present amino
CC acid sequence represents a human CGSD-associated protein of the
CC invention.
XX
XX
SQ Sequence 68 AA:

Alignment Scores:
Pred. No.: 76.7 Length: 68
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 8 Gaps: 0

US-10-030-271-3 (1-1883) x ADN01127 (1-68)

OY 366 GCTCTGCCGACATGCCGCGGCGCTC 340
|||
18 ALaLeuAlaLeuProAlaAlaLeu 26

RESULT 37
AAU41429
ID AAU41429 standard; protein; 71 AA.
AC AAU41429;
XX
XX 13-FEB-2002 (first entry)
XX
XX
XX Propionibacterium acnes immunogenic protein #2325.
XX
XX
XX SAPHO syndrome; myovitis; acne; pustulosis; hyperostosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX
XX Propionibacterium acnes.
OS
XX
XX WO200181581-A2.
XX
XX
XX 01-NOV-2001.
XX
XX
XX 20-APR-2001; 2001WO-US012865.
XX
XX
XX 21-APR-2000; 2000US-0199047P.
XX
XX 02-JUN-2000; 2000US-0208841P.
XX
XX 07-JUL-2000; 2000US-0216747P.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhacia A;
PI L'maisonneuve J, Zhang Y, Jen S Carter D;

XX	WPI; 2001-616774/71.
DR	N-PSDB; AAS59515.
XX	
PT	Propionibacterium acnes polypeptides and nucleic acids useful for
PT	vaccinating against and diagnosing infections, especially useful for
PT	treating acne vulgaris.
XX	
PS	Example 1; SEQ ID NO 2624; 1069pp; English.
XX	
CC	Sequences AAU9105-AU68017 represent Propionibacterium acnes immunogenic
CC	polypeptides. The proteins and their associated DNA sequences are used in
CC	the treatment, prevention and diagnosis of medical conditions caused by
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC	pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC	P. acnes is also involved in infections of bone, joints and the central
CC	nervous system, however it is particularly involved in the inflammatory
CC	lesions associated with acne vulgaris. A method for detecting the
CC	presence or absence of P. acnes in a patient comprises contacting a
CC	sample with a binding agent that binds to the proteins of the invention
CC	and determining the amount of bound protein in the sample. The
CC	polypeptides may be used as antigens in the production of antibodies
CC	specific for P. acnes proteins. These antibodies can be used to
CC	downregulate expression and activity of P. acnes polypeptides and
CC	therefore treat P. acnes infections. The antibodies may also be used as
CC	diagnostic agents for determining P. acnes presence, for example, by
CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC	this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 71 AA;
	Alignment Scores:
	Pred. No.: 76.2 Length: 71
	Score: 9.00 Matches: 9
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 1.46% Indels: 0
	DB: 4 Gaps: 0
	US-10-030-271-3 (1-1883) x AAU41429 (1-71)
OY	159 CCAGCAGGGGCGGGGTGCACCGGA 133
Db	6 ProAlaArgGlyArgGlyArgProGly 14
RESULT 38	
ID	ABM37948 standard; protein; 71 AA.
XX	
AC	ABM37948;
XX	
DT	20-OCT-2003 (first entry)
XX	
DE	Propionibacterium acnes predicted ORF-encoded polypeptide #2624.
XX	
KW	Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX	immunostimulant; immune response; vaccine.
OS	Propionibacterium acnes.
PN	WO2003033515-A1.
PD	24-APR-2003.
XX	
PF	11-OCT-2002; 2002MO-US032727.
XX	
PR	15-OCT-2001; 2001US-0097825.
PA	(CORI-) CORIXA CORP.
XX	
MI	Mitcham JV, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JV;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vailleve-Douglas J;
DR WPI; 2003-381789/36.
XX N-PSDB; ACF64444.
PS Example 1; SEQ ID NO 2624; 1481pp; English.
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM55624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 71 AA;
Alignment Scores:
Pred. No.: 76.2 Length: 71
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 0
US-10-030-271-3 (1-1883) x ABM57948 (1-71)
QY 159 CCAGCAGGGGCGGGGTCGACCCGGA 133
Db 6 ProXlaArgGlyArgGlyArgProGly 14
RESULT 39
ID AAV62949 standard; protein; 89 AA.
XX AAV62949;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #23845.
XX
KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopthic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX

PN W0200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX (CORI-) CORIXA CORP.
XX
PI Szeily YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
XX N-PSDB; AAS59630.
XX
PS Example 1; SEQ ID NO 24144; 1069pp; English.
XX
CC Sequences AAV39105-AAV68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 89 AA;
Alignment Scores:
Pred. No.: 73.8 Length: 89
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
Gaps: 0
US-10-030-271-3 (1-1883) x AAV62949 (1-89)
QY 350 CCGCGGCGCTCCAGCTCCAGCAGAGC 324
Db 66 ProXlaArgSerSerSerSerArgSer 74
RESULT 40
ID AAV59468 standard; protein; 89 AA.
XX AAV59468;
XX
AC AAV59468;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #24144.
XX
KM Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KM immunostimulant; immune response; vaccine.
XX

XX OS Propionibacterium acnes.
XX PN WO2003033515-A1.
XX PD 24-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032727.
XX PR 15-OCT-2001; 2001US-00978825.
XX PA (CORI-) CORIYA CORP.
XX PI Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL,
XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX PI Barth B, Vallieue-Douglas J;
XX DR WPI; 2003-381789/36.
XX DR N-PSDB; ACP64559.
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX PS Example 1; SEQ ID NO 24144; 1481pp; English.
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention; antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
XX CC and the kit is useful for performing a diagnostic assay. The present
XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
XX CC reading frame) contained within the P. acnes polynucleotides of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at http://wipo.int/pub/published_pct_sequences
XX SQ Sequence 89 AA;
XX
XX Alignment Scores:
XX Pred. No.: 73.8
XX Score: 9.00
XX Percent Similarity: 100.00%
XX Best Local Similarity: 100.00%
XX Query Match: 1.46%
XX DB: 6
XX Length: 89
XX Matches: 9
XX Conservative: 0
XX Mismatches: 0
XX Indels: 0
XX Gaps: 0
XX
XX US-10-030-271-3 (1-1883) x ABM59468 (1-89)
XX
XX QY 350 CCGCGGCGCTCCAGCTCCAGCAGAGC 324
XX DB 66 ProkArgArgSerSerSerSerArgSer 74
XX
XX RESULT 41
XX ABO81797
XX ID ABO81797 standard; protein; 91 AA.

XX AC ABO81797;
XX XX 29-JUL-2004 (first entry)
XX DT
XX DE Pseudomonas aeruginosa polypeptide #13972.
XX DE
XX KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX OS Pseudomonas aeruginosa.
XX XX US6551795-B1.
XX PN
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX XX
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX DR WPI; 2003-615309/58.
XX DR N-PSDB; ABD15368.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 30543; 455pp; English.
XX XX
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biochip technology. Sequences ABO67826-
XX CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX CC sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 91 AA;
XX
XX Alignment Scores:
XX Pred. No.: 73.5
XX Score: 9.00
XX Percent Similarity: 100.00%
XX Best Local Similarity: 100.00%
XX Query Match: 1.45%
XX DB: 7
XX Length: 91
XX Matches: 9
XX Conservative: 0
XX Mismatches: 0
XX Indels: 0
XX Gaps: 0
XX
XX US-10-030-271-3 (1-1883) x ABO81797 (1-91)
XX
XX QY 628 AGACGGCGGCGAGAGGCGCCAGCC 654
XX DB 5 ArgArgArgArgGlyAlaProAla 13
XX
XX RESULT 42
XX AAU46695
XX ID AAU46695 standard; protein; 125 AA.
XX AC AAU46695;
XX XX
XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #7591.
 XX XX
 XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.
 XX OS
 XX Propionibacterium acnes.
 XX PN
 XX WO200181581-A2.
 XX PD
 XX 01-NOV-2001.
 XX PF
 XX 20-APR-2001; 2001WO-US012865.
 XX PR
 XX 21-APR-2000; 2000US-0199047P.
 XX PR 02-JUN-2000; 2000US-0208841P.
 XX PR 07-JUL-2000; 2000US-0216747P.
 XX PA
 XX (CORI-) CORIXA CORP.
 XX PI
 XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX DR
 XX WPI; 2001-616774/71.
 XX DR
 XX N-PSDB; AAS59534.
 XX PT
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX PS
 XX Example 1; SEQ ID NO 7890; 1069pp; English.
 XX XX
 CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting the
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ
 XX Sequence 125 AA;
 Alignment Scores:
 Pred. No.: 70.2 Length: 125
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 4 Gaps: 0
 US-10-030-271-3 (1-1883) x AAU46695 (1-125)
 QY 656 GCGGCTGGGGCCCTCTCCGCGCCGCT 630
 Db 15 AlahlaaglyahlaProleuArgArg 23
 RESULT 43
 ABM43214

ID ABM43214 standard; protein; 125 AA.
 XX AC
 XX ABM43214;
 XX XX
 DT 20-OCT-2003 (first entry)
 XX XX
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #7890.
 XX KM
 XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KM immunostimulant; immune response; vaccine.
 XX OS
 XX Propionibacterium acnes.
 XX PN
 XX WO2003033515-A1.
 XX PD
 XX 24-APR-2003.
 XX PF
 XX 11-OCT-2002; 2002WO-US032727.
 XX PR
 XX 15-OCT-2001; 2001US-00978825.
 XX PA
 XX (CORI-) CORIXA CORP.
 XX PI
 XX Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 XX PI
 XX Barth B, Vallieve-Douglas J;
 XX XX
 XX WPI; 2003-381789/36.
 XX DR
 XX N-PSDB; ACF64463.
 XX PT
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX PS
 XX Example 1; SEQ ID NO 7890; 1481pp; English.
 XX XX
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ
 XX Sequence 125 AA;
 Alignment Scores:
 Pred. No.: 70.2 Length: 125
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 6 Gaps: 0

US-10-030-271-3 (1-1883) x ABM43214 (1-125)

QY 656 GCGGCTGGGGCCCTCTCCGCCCGCT 630
 DB 15 AIAAIAcIyAlaPLeuArgArgArg 23

RESULT 44
 ABO70723
 ID ABO70723 standard; protein; 148 AA.
 XX
 AC ABO70723;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #2898.
 XX
 KM Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 PD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfeld MJ, Nolling J, Deloughery C, Bush D;
 XX
 DR WPI; 2003-615309/58.
 DR N-PSDB; ABD04294.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 19469; 455bp; English.
 XX
 CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biotechnology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 XX
 SQ Sequence 148 AA;

Alignment Scores:
 Ptd. No.: 68.5 Length: 148
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.46% Indels: 0
 DB: 7 Gaps: 0

QY 1166 CGTCCTCCAGAGAGGTG3CCCGGA 1140

DB 104 ArgProProArgArgArgTrProGly 112
 RESULT 45
 AAE03094
 ID AAE03094 standard; protein; 211 AA.
 XX
 AC AAE03094;
 XX
 DT 10-AUG-2001 (first entry)
 XX
 DE Human gene 8 encoded secreted protein HATM46. SEQ ID NO:128.
 XX
 KM Human; secreted protein; proliferative disorder; cancer; tumour;
 KM foetal abnormality; developmental abnormality; haematopoietic disorder;
 KM immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KM inflammation; allergy; neurological disorder; Alzheimer's disease;
 KM Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KM skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KM cardiovascular disorder; angiodysplasia; kidney disorder;
 KM gastrointestinal disorder; pregnancy-related disorder; gene therapy;
 KM endocrine disorder; infection; wound healing; veterinary; cell culture;
 KM chemotaxis; food additive; binding partner identification.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..1 /label= Signal_peptide
 FT Protein 2..211 /note= "Mature secreted protein"
 FT MISC-difference 64 /label= Unknown
 FT /note= "Encoded by GSN"
 FT MISC-difference 113 /label= Unknown
 FT /note= "Encoded by GNT"
 XX
 PN W0200132676-A1.
 XX
 PD 10-MAY-2001.
 XX
 PF 25-OCT-2000; 2000MO-US029365.
 XX
 PR 29-OCT-1999; 99US-0162237P.
 PR 21-JUL-2000; 2000US-0219666P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Komatsoulis GA, Shi Y, Olsen HS, Soppet DR;
 XX
 DR WPI; 2001-328773/34.
 DR N-PSDB; AAD07613.
 XX
 PT Nucleic acids encoding 25 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating e.g. Gaucher's disease,
 PT Alzheimer's disease, Schmitzer syndrome, Creutzfeldt-Jacob disease,
 PT diabetes mellitus and multiple sclerosis.
 XX
 PS Claim 11; Page 480-481; 546bp; English.
 XX
 CC AAD07571-AAD07645 represent CDNs corresponding to 25 human secreted
 CC protein genes, and AAE03052-AAE03126 represent the proteins they encode.
 CC AAE03127-AAE03150 represent human secreted protein fragments. The genes
 CC and their corresponding secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 25 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,

CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein of the invention
XX
SQ Sequence 211 AA;

Alignment Scores:
Pred. No.: 65 Length: 211
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 0
US-10-030-271-3 (1-1883) x AAE03094 (1-211)

QY 811 CCAGCTGCCGCGCAGCGCTGGGGCC 785
Db 126 ProAlaAlaProAlaProAla 134

RESULT 46
ABG63622
ID ABG63622 standard; protein; 211 AA.
XX
AC ABG63622;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human albumin fusion protein #297.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW hematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfectivity; antiinflammatory; antitumor;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiParkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
XX Synthetic.
XX
PN WO200177137-A1.
XX
PD 18-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US011988.
XX
XX 12-APR-2000; 2000US-0229358P.
PR 25-APR-2000; 2000US-0199384P.
PR 21-DEC-2000; 2000US-0256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2002-010886/01.
XX
PT New fusion protein for treating disease e.g. diabetes comprises an

PT albumin fused to a therapeutic protein.
XX
PS Claim 1; Page 758; 2102dp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
CC therapeutic protein X and human albumin (HA, also known as human serum
CC albumin, HSA). The proteins are useful for treating a disease or disorder
CC that may be modulated by therapeutic protein X. The albumin extends the
CC shelf-life of protein X, and may increase its biological in vitro/in vivo
CC activity. The protein is useful for treating and diagnosing disorders
CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
CC disease, ulcerative colitis), immune disorders (e.g. acquired
CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
XX
SQ Sequence 211 AA;

Alignment Scores:
Pred. No.: 65 Length: 211
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: Gaps: 0
US-10-030-271-3 (1-1883) x ABG63622 (1-211)

QY 811 CCAGCTGCCGCGCAGCGCTGGGGCC 785
Db 126 ProAlaAlaProAlaProAla 134

RESULT 47
ADA57327
ID ADA57327 standard; protein; 211 AA.
XX
AC ADA57327;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human secreted protein #39.
XX
KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
KW cytostatic; cerebroprotective; neuroprotective; nootropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology.
XX
XX
OS Homo sapiens.
XX
XX
PN WO2002102994-A2.
XX
PD 27-DEC-2002.
XX
PF 19-MAR-2002; 2002WO-US008278.
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-167512/16.
XX
XX N-PSDB; ADA56431.
XX

US-10-030-271-3 (1-1883) x ADJ72183 (1-262)

Qy	930	CGCCCTGCTGCAGGCCCTGCGGGGGGT	956
Db	18	ArgProAlaAlaGlyProAlaGlyArg	26

Search completed: March 22, 2005, 03:00:50
Job time : 279.678 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 22, 2005, 03:04:23 ; Search time 498.739 Seconds
(without alignments)
2495.932 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 620
Sequence: 1 aggcgcataatagagaagg.....ctggaaaggcactagctggg 1883

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1401741 segs, 330541175 residues

Word size: 1

Total number of hits satisfying chosen parameters: 268658

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=x1p
-G/cgn2_1/USPRO.spool.p/US10030271/runat_21032005_153550_26862/app_query_fastra_1.3150
-DB=Published Applications_AA -QFMT=fastran -SURFIX=oligo.rapb -MINMATCH=0.1
-LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.ccl -LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10030271.0CGN_1_1_439 @runat_21032005_153550_26862
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-Fgapop=6 -Fgapext=7 -Ygapop=60 -Ygapext=60 -Delop=6 -Delext=7

Database : Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result

Query

No.	Score	Match Length	DB	ID	Description
1	257	41.5	366	14 US-10-106-698-4626	Sequence 4626, Ap
2	247	39.8	304	14 US-10-013-477-11	Sequence 11, Appl
3	247	39.8	326	14 US-10-296-539-1	Sequence 1, Appl
4	239	38.5	318	14 US-10-001-254-18	Sequence 18, Appl
5	204	32.9	319	9 US-09-925-302-758	Sequence 758, App
6	204	32.9	319	10 US-09-925-302-758	Sequence 758, App
7	157	25.3	242	15 US-10-296-115-1220	Sequence 1220, Ap
8	118	19.0	217	9 US-09-799-777-26	Sequence 28, Appl
9	67	10.8	101	14 US-10-001-254-8	Sequence 8, Appl
10	32	5.2	40	14 US-10-013-477-20	Sequence 20, Appl
11	13	2.1	210	9 US-09-733-167-6	Sequence 6, Appl
12	13	2.1	318	9 US-09-935-223-4	Sequence 4, Appl
13	13	2.1	318	9 US-09-935-223-6	Sequence 6, Appl
14	13	2.1	318	9 US-09-733-167-1	Sequence 1, Appl
15	13	2.1	318	9 US-09-733-167-3	Sequence 3, Appl
16	11	1.8	107	16 US-10-437-963-107955	Sequence 107955,
17	11	1.8	145	16 US-10-437-963-166242	Sequence 166242,
18	10	1.6	146	16 US-10-001-254-38	Sequence 38, Appl
19	9	1.5	66	15 US-10-424-599-231701	Sequence 231701,
20	9	1.5	90	16 US-10-767-701-47327	Sequence 47327, A
21	9	1.5	94	16 US-10-767-701-55826	Sequence 55826, A
22	9	1.5	115	16 US-10-424-599-225144	Sequence 225144,
23	9	1.5	116	16 US-10-767-701-48342	Sequence 48342, A
24	9	1.5	128	16 US-10-437-963-142899	Sequence 142899,
25	9	1.5	131	16 US-10-437-963-142899	Sequence 142899,
26	9	1.5	133	16 US-10-437-963-164498	Sequence 164498,
27	9	1.5	172	16 US-10-437-963-204745	Sequence 204745,
28	9	1.5	211	11 US-09-833-245-369	Sequence 369, App
29	9	1.5	215	16 US-10-437-963-124217	Sequence 124217,
30	9	1.5	216	16 US-10-425-114-56323	Sequence 56323, A
31	9	1.5	219	16 US-10-437-963-197739	Sequence 197739,
32	9	1.5	243	16 US-10-437-963-175466	Sequence 175466,
33	9	1.5	262	17 US-10-211-028-31	Sequence 31, Appl
34	9	1.5	268	15 US-10-425-114-60272	Sequence 60272, A
35	9	1.5	330	16 US-10-437-963-104157	Sequence 104157,
36	9	1.5	336	16 US-10-781-014-490	Sequence 490, App
37	9	1.5	370	14 US-10-285-976-1	Sequence 1, Appl
38	9	1.5	370	15 US-10-302-812-50	Sequence 50, Appl
39	9	1.5	370	17 US-10-847-972-2	Sequence 2, Appl
40	9	1.5	335	9 US-09-738-626-6644	Sequence 6644, Ap
41	9	1.5	388	14 US-10-156-761-13365	Sequence 13365, A
42	9	1.5	410	15 US-10-425-114-38092	Sequence 38092, A
43	9	1.5	456	15 US-10-369-499-20883	Sequence 20883, A
44	9	1.5	477	16 US-10-437-963-155657	Sequence 155657,
45	9	1.5	478	14 US-10-156-761-10964	Sequence 10964, A
46	9	1.5	503	8 US-08-459-455-36	Sequence 36, Appl
47	9	1.5	505	9 US-09-888-243-5	Sequence 5, Appl
48	9	1.5	517	15 US-10-425-114-62674	Sequence 62674, A
49	9	1.5	564	15 US-10-425-114-42986	Sequence 42986, A
50	9	1.5	614	15 US-10-282-122A-76690	Sequence 76690, A
51	9	1.5	675	16 US-10-437-963-149484	Sequence 149484,
52	9	1.5	901	16 US-10-437-963-142715	Sequence 142715,
53	9	1.5	955	16 US-10-408-765A-1972	Sequence 1972, Ap
54	9	1.5	962	15 US-10-276-77-2346	Sequence 2346, Ap
55	9	1.5	2031	15 US-10-282-122A-62686	Sequence 62686, A
56	9	1.5	10	10 US-09-572-404B-1528	Sequence 1528, Ap
57	9	1.5	10	10 US-09-572-404B-1530	Sequence 1530, Ap
58	8	1.3	15	9 US-09-879-957-150	Sequence 150, App
59	8	1.3	15	14 US-10-185-050-94	Sequence 94, Appl
60	8	1.3	15	14 US-10-161-791-307	Sequence 307, App
61	8	1.3	32	15 US-10-430-011-115	Sequence 115, App
62	8	1.3	34	16 US-10-437-963-186712	Sequence 186712,
63	8	1.3	41	15 US-10-424-599-176515	Sequence 176515,
64	8	1.3	43	9 US-09-896-888A-27	Sequence 27, Appl
65	8	1.3	43	16 US-10-437-963-102706	Sequence 102706,
66	8	1.3	45	15 US-10-264-049-4309	Sequence 4309, Ap

FILE REFERENCE: PT002P1
CURRENT APPLICATION NUMBER: US/10/013,477
CURRENT FILING DATE: 2001-12-13
PRIOR APPLICATION NUMBER: 09/669,445
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: PCT/US00/06642
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: 60/126,018
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/139,638
PRIOR FILING DATE: 1999-06-17
PRIOR APPLICATION NUMBER: 60/149,449
PRIOR FILING DATE: 1999-08-18
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 11
LENGTH: 304
TYPE: PRT
ORGANISM: Homo sapiens
US-10-013-477-11

Alignment Scores:

Pred. No.:	1,166-202	Length:	304
Score:	247.00	Matches:	247
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	39.84%	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-013-477-11 (1-304)

QY 361 GAGAGCAACTGCGGCTGCTGGGCAACTCTGCGCGTGTGCGCCGCAAGCACTGCTG 420
DB 58 GUSerAsnLeuArgLeuGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 77
QY 421 CGGACCTGGCGGCGCAAGCGCGCGCGCGCGCGCTCTCCAGAGCGCTATAGCTATGGCACC 480
DB 78 ProHisLeuAlaArgHisArgArgArgProValSerProGluArgTyrSerTyrGlyThr 97
QY 481 TCCAGCTCTTCAAG 540
DB 98 Ser 117
QY 541 GCAATTTCTCAGAGAGGTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 118 AlaAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 137
QY 601 AGTGGGCG 660
DB 138 SerArgGlyArgProSerGlyGlyAlaArgArgArgArgArgArgArgArgArgArg 157
QY 661 CAGCAGCAGTCAAGCG 720
DB 158 GlnGlnGlnSerGlnProAlaArgProSerSerGlnGlnGlnGlnGlnGlnGlnGln 177
QY 721 CTCGGGTTTCAGAGAGTACTGCGAGCATGGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
DB 178 LeuArgValArgHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
QY 781 CGGCG 840
DB 198 ArgArgProGlnAlaLeuAlaArgGlnLeuAspValPheGlyGlnAlaThrAlaValLeu 217
QY 841 CGGTCAAGGAGACTGGGCTCTGTGTTTGTGACATCAAGTTTCAAGCTCTCTATCTG 900
DB 218 ArgSerArgAspLeuGlnSerValValCysAspIleLeuSerSerGlnLeuSerTyrLeu 237
QY 901 GAGCGCTTCTGAGGCGACTGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
DB 238 AspAlaPheTrpGlyAspTyrLeuSerGlyAlaLeuGlnGlnAlaLeuArgGlyValPhe 257
QY 961 CTGACTGAGGCGCTGCGAGAGGCTGTGGCGCGGAGGCTGTGCTGCTGCTGCTGCTG 1020

DB 258 LeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeuValSerVal 277
QY 1021 GATGAGCTGACTATAGAGGTCGCGCGCGCGCGCGCGCTGCTGCTGATGAGAGAGAGAG 1080
DB 278 AspGlnAlaAspTyrGlnAlaGlyArgArgArgLeuLeuLeuMetGlnGlnGlnGly 297
QY 1081 CGGCG 1101
DB 298 ArgArgProThrGlnAlaSer 304

RESULT 3

US-10-296-539-1
Sequence 1, Application US/10296539
Publication No. US20030165933A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: TANG, Y. Tom
APPLICANT: AZIMZAI, Yalda
APPLICANT: YOE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: DING, Li
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: PATTERSON, Chandra
APPLICANT: BAUGHN, Mariah R.
TITLE OF INVENTION: REGULATORS OF APOPTOSIS
FILE REFERENCE: PI-0307 PCT
CURRENT APPLICATION NUMBER: US/10/296,539
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: 60/209,407; 60/250,326
PRIOR FILING DATE: 2000-06-01; 2000-11-30
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 326
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030165933A1 3102521CD1
US-10-296-539-1

Alignment Scores:

Pred. No.:	1,146-202	Length:	326
Score:	247.00	Matches:	247
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	39.84%	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-296-539-1 (1-326)

QY 361 GAGAGCAACTGCGGCTGCTGGGCAACTCTGCGCGTGTGCGCCGCGCAAGCACTGCTG 420
DB 80 GUSerAsnLeuArgLeuGlnLeuLeuArgValLeuAlaArgHisAspLeuLeu 99
QY 421 CGGACCTGGCGGCGCAAGCGCGCGCGCGCGCGCTCTCCAGAGCGCTATAGCTATGGCACC 480
DB 100 ProHisLeuAlaArgHisArgArgArgProValSerProGluArgTyrSerTyrGlyThr 119
QY 481 TCCAGCTCTTCAAG 540
DB 120 SerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 139
QY 541 GCAATTTCTCAGAGAGGTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 140 AlaAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 159
QY 601 AGTGGGCG 660
DB 160 SerArgGlyArgProSerGlyGlyAlaArgArgArgArgArgArgArgArgArgArg 179
QY 661 CAGCAGCAGTCAAGCG 720

Db 180 GINGINGInserGIuProAlaArgProSerSerGIuGIyLyValThrCyAspIleArg 199
Qy 721 CTCGGGTTGAGACAGACTGCTGCGAGCATGGGCCAGCTTGAGAGGCGCTGGCATCC 780
Db 200 LeuArgValArgAlaGIuTrCySerGluHISGlyProAlaLeuGIuInGIyAlaAlaSer 219
Qy 781 CGGCGGCCCCAGGCGCTGGCGCGGCGAGCTGGAACGCTTTGGGCGAGGCCACCGCAGTGGCTG 840
Db 220 ArgArgProGIuAlaLeuAlaArgGIuLeuAspValPheGIyGIuAlaThrAlaValLeu 239
Qy 841 CGCTCAAGGAGCCTGGGCTCTGTGGTTTGTGACATCAATGTTCTCAGAGCTCTCTATCTG 900
Db 240 ArgSerArgAspLeuGIySerValValCyAspIleIySphSerGIuLeuSerIyLeu 259
Qy 260 AspAlaPheTrpGIyAspIyTrLeuSerGIyAlaLeuLeuGIuAlaLeuArgIyValPhe 279
Db 961 CTGACTGAGGCGCTGCGAAGGCTGTGGGCGCGGAGGCTGTTCGCTGTGTCAGTGTG 1020
Qy 280 LeuTrnGIuAlaLeuArgIuAlaValGIyArgGIuAlaValArgLeuValSerVal 299
Db 1021 GATGAGCTGACTATGAGGCTGGCGCGCGCTGTGTGATGAGAGGAGGAGGCGG 1080
Qy 300 AspGIuAlaAspTrpGIuAlaGIyArgArgArgLeuLeuMetGIuInGIyGIy 319
Db 1081 CGGCGCGCGAGAGGCGCTCC 1101
Qy 320 ArgArgProThrGIuAlaSer 326

RESULT 4

US-10-001-254-18
; Sequence 18, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sug Hyung
; APPLICANT: Roth, Wilfred
; APPLICANT: Steiner-liem, Frank
; TITLE OF INVENTION: No. US20030049702A1el Death Domain Proteins
; FILE REFERENCE: F-L0 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-254-18

Alignment Scores:

Pred. No.: 8,33e-196 Length: 318
Score: 239.00 Matches: 239
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.55% Indels: 0
DB: 14 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-001-254-18 (1-318)

Qy 361 GAGAGCAACCTGGGCTGTGGGGCAACTCTGGCGGTGTGGCCCGCAGACTGTCTG 420
Db 80 GluSerAsnLeuArgLeuLeuGIyInLeuLeuArgValLeuAlaTrnGHSAspLeuLeu 99
Qy 421 CGGCACTGGCGCGCAGGCGCGCGCCAGTGTCTCCAGAACGCTATAGCTATGGCACC 480

Db 100 ProHISLeuAlaArgIyAspArgArgProValSerProGIuArgTrISerIyGIyTr 119
Qy 481 TCCAGCTTTCAAGAGAGCAGAGGATAGCTCCGTCGCGCTGGCAGTCAAGACTTCT 540
Db 120 SerSerSerSerIyArgTrnGIuGIySerCyAspArgArgArgInSerSerSer 139
Qy 541 GCMAATTTCAAGAGGTTAGTGGAGACAGGCTCCCCCAACCAAGGCGAGCGCGCG 600
Db 140 AlaAsnSerGIuInGIyGIuTrnTrnGIySerProTrnIyAspArgIuAlaArg 159
Qy 601 AGTCGGGCGCGCGCAGTGTGTGTGCAGACGCGCGGAGAGGAGGCGCGCAGCC 660
Db 160 SerArgGIyArgProSerGIyAlaArgAlaArgArgGIyAlaProAlaAlaPro 179
Qy 661 CAGACAGCTCAGAGCGCGCAGACCTTCTGTGAAGCAAGTGAACCTGTGACATCCG 720
Db 180 GINGINGInserGIuProAlaArgProSerSerGIuGIyLyValThrCyAspIleArg 199
Qy 721 CTCGGGTTGAGACAGACTGCTGCGAGCATGGGCCAGCTTGAGAGGCGCTGGCATCC 780
Db 200 LeuArgValArgAlaGIuTrCySerGluHISGlyProAlaLeuGIuInGIyAlaAlaSer 219
Qy 781 CGGCGGCCCCAGGCGCTGGCGCGGCGAGCTGACATCAATGTTCTCAGAGCTCTCTATCTG 840
Db 220 ArgArgProGIuAlaLeuAlaArgGIuLeuAspValPheGIyGIuAlaThrAlaValLeu 239
Qy 240 ArgSerArgAspLeuGIySerValValCyAspIleIySphSerGIuLeuSerIyLeu 259
Db 841 CGCTCAAGGAGCCTGGGCTCTGTGGTTTGTGACATCAATGTTCTCAGAGCTCTCTATCTG 900
Qy 901 GAGCGCTTTGGGCGGCACTCACTGAGTGGCGCGCTGCTGACAGCCCTGGGCGCGTTC 960
Db 260 AspAlaPheTrpGIyAspIyTrLeuSerGIyAlaLeuLeuGIuAlaLeuArgIyValPhe 279
Qy 961 CTGACTGAGGCGCTGCGAAGGCTGTGGGCGCGGAGGCTGTTCGCTGTGTCAGTGTG 1020
Db 280 LeuTrnGIuAlaLeuArgIuAlaValGIyArgGIuAlaValArgLeuValSerVal 299
Qy 1021 GATGAGCTGACTATGAGGCTGGCGCGCGCTGTGTGATGAGAGGAGGAGG 1077
Db 300 AspGIuAlaAspTrpGIuAlaGIyArgArgArgLeuLeuMetGIuInGIyGIy 318

RESULT 5

US-09-925-302-758
; Sequence 758, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 758
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-758

Alignment Scores:

Pred. No.: 8,54e-166 Length: 319
Score: 204.00 Matches: 304
Percent Similarity: 99.67% Conservative: 0
Best Local Similarity: 99.67% Mismatches: 1
Query Match: 32.90% Indels: 0
DB: 9 Gaps: 0

QY 748 CATGGCCAGCCTTTGAGAGAGGGCGTGGATCCCGGGGGCCCGAGGGCTGGCGCCGAG 807
DB 249 HsGlyProAlaLeuGluGlnGlyValAlaSerArgProGlnAlaLeuAlaArgGln 268
QY 808 CTGAGCGTGTGGAGGCGCCAGCGAGTGTGGCTCAAGGAGCACTGGCTGTGGTT 867
DB 269 LeuAspValPheGlyGlnAlaThrAlaValLeuArgSerArgAspLeuGlySerVal 288
QY 868 TGTGACATCAAGTTCTCAGAGCTCTCCTATCTGGACGCGCTTGGGGCGACTACCTGAGT 927
DB 289 CysAspLeuSerPheSerGluLeuSerTyrLeuAspAlaPheTyrGlyAspTyrLeuSer 308
QY 928 GGGGCGCTGCTGCAG 942
DB 309 GlyAlaLeuLeuGln 313
RESULT 7
US-10-296-115-1220
; Sequence 1220, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1220
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1220
Alignment Scores:
Pred. No.: 1.8e-125 Length: 242
Score: 157.00 Matches: 157
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.32% Indels: 0
DB: Gaps: 0
US-10-030-271-3 (1-1883) x US-10-296-115-1220 (1-242)
QY 631 CGGGGGGAGAGGGGGCCGAGCCCGAGAGAGTCAGAGCCCGGAGACTTCC 690
DB 86 ArgArgArgArgGlyAlaProAlaAlaProGlnGlnSerGluProAlaArgProSer 105
QY 691 TCTGAAGCAAAAGTACCTGTGACATCCGGTCCGGGTTGAGAGAGTACTGGAGCAT 750
DB 106 SerGluGlyValThrCysAspLeuArgValAlaArgAlaGluTyrCysGluHis 125
QY 751 GGGCGACCTTGGAGAGGGGGTGGATCCCGGGCGCCCGAGGGCGTGGCGCGCACTG 810
DB 126 GlyProAlaLeuGluGlnGlyValAlaSerArgProGlnAlaLeuAlaArgGlnLeu 145
QY 811 GAGCTGTTGGGGAGGCGCACCGAGTGTGGCTCAAGGGGAGCGGCTGTGGTTGT 870
DB 146 AspAlaPheGlyGlnAlaThrAlaValLeuArgSerArgAspLeuGlySerValCys 165
QY 871 GACATCAAGTTCTCAGAGCTCTCCTATCTGGACGCGCTTGGGGCGACTACCTGAGTGC 930
DB 166 AspLeuSerPheSerGluLeuSerTyrLeuAspAlaPheTyrGlyAspTyrLeuSerGly 185
QY 931 GCGCTGTGAGAGGCTGCGGGGCGTGTCTGACTGAGGCGCTGCGAGAGCTGTGGGC 990
DB 186 AlaLeuGlnAlaLeuAlaArgGlyValPheLeuThrGlnAlaLeuAlaValGly 205
QY 991 CGGAGAGCTGTGGCTGTGGTGTGATGTGATAGGCTGACTATGAGGCTGGCGGGCGC 1050

DB 206 ArgGluAlaValArgLeuLeuValSerValAspGluAlaAspTyrGluAlaArgArg 225
QY 1051 CGCCTGTGCTGATGAGAGAGAGGGGGCGCGCCCGGAGAGGCTTCC 1101
DB 226 ArgLeuLeuLeuMetGluGluGluGlyArgArgProThrGluAlaSer 242
RESULT 8
US-09-799-777-26
; Sequence 26, Application US/09799777
; Patent No. US20020091244A1
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; Hillman, Jennifer L.
; Cortley, Neil C.
; Guegler, Karl J.
; Baugh, Mariah
; Sather, Susan
; Shah, Puri
; TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS
; NUMBER OF SEQUENCES: 154
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/799,777
; FILING DATE: 06-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/002,485
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BILKINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0459 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: UTRSN0706
; CLONE: 1638407
; SEQUENCE DESCRIPTION: SEQ ID NO: 26 :
US-09-799-777-26
Alignment Scores:
Pred. No.: 5.08e-92 Length: 217
Score: 118.00 Matches: 118
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.03% Indels: 0
DB: Gaps: 0
US-10-030-271-3 (1-1883) x US-09-799-777-26 (1-217)
QY 712 GACATCGGCTCCGGGTTTCAGAGAGTACTGCGAGCATGGCCAGCTTTGAGAGAGGC 771
DB 88 AspLeuArgLeuArgValArgAlaGluTyrCysGluHisGlyProAlaLeuGluGlnGly 107
QY 772 GTGGATCCGGGAGGCGCCCGAGGCGCTGGCGGAGCTGAGAGCTTTGGGAGGCAACC 831

```
Db 108 ValAlaSerArgGlyProGlnAlaLeuAlaArgGlnLeuSerValIleGlyGlnAlaThr 127
Qy 832 GCAGTGTGGCGCTCAGAGGACCTGGGCTCTGGTTGTGACATCAAGTTCTAGAGCTC 891
Db 128 AlValLeuArgSerArgAspLeuGlySerValValCysAspIleLysPheSerGlnLeu 147
Qy 892 TCCATCTGGAAGCGCTTCTGGGGGCGACTACCTGAGTGCGCGCCCTGCTGACAGCCCTGGG 951
Db 148 SerTyrLeuSerAlaPheTrpGlyAspTyrLeuSerGlyAlaLeuGlnAlaLeuArg 167
Qy 952 GGGGTGTCTCTGATGAGAGCCCTCGAGAGGCTGTGGCCGGGAGGCTGTTCGCTGTG 1011
Db 168 GlyValPheLeuThrGlnAlaLeuArgGlnAlaValGlyArgGlnAlaValArgLeuLeu 187
Qy 1012 GTGAGTGTGAGTGAAGCTGACTATGAGGCTGGCGGCGCGCTGTGCTGATG 1065
Db 188 ValSerValAspGlnAlaAspTyrGlnAlaGlyArgArgLeuLeuMet 205
```

```
RESULT 9
US-10-001-254-8
; Sequence 8, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Koch, Wilfried
; APPLICANT: Stehner-Blewen, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-LJ 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-254-8
```

```
Alignment Scores:
Pred. No.: 3,16e-48 Length: 101
Score: 67.00 Matches: 67
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.81% Indels: 0
DB: Gaps: 0
```

US-10-030-271-3 (1-1883) x US-10-001-254-8 (1-101)

```
Qy 157 TGGAGAGAGATGATGCTGCTGACTACTACGGGATGCTGTGCTTACCGTATGTTGAG 216
Db 1 TrpIleTlnuAspGlnCysLeuAspTyrTylGlyMetLeuSerLeuHISArgMetPheGln 20
Qy 217 GTGGTGGGGGCGCAATGACCGAGTGCAGGCTGAGCTCCGCGCTTCTGCTGATGAG 276
Db 21 ValValIleTylGlnLeuThrGlnCysGlnLeuGlnLeuValAlaPheLeuLeuAspGln 40
Qy 277 GCTCTGTCGCGCGCGGAGGCTTAGCCCGGCGCGGCGGCTTAGAGCTTCTGCTGAG 336
Db 41 AlaProGlyAlaIleGlyGlyLeuAlaArgAlaArgSerGlyLeuGlnLeuLeuGln 60
Qy 337 CTGAGAGCGCGCGGCGGAGCTGC 357
Db 61 LeuGlnArgArgGlyGlnCys 67
```

```
RESULT 10
US-10-013-477-20
; Sequence 20, Application US/10013477
; Publication No. US20030049732A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Apoptosis Related Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO02P1
; CURRENT APPLICATION NUMBER: US/10/013,477
; CURRENT FILING DATE: 2001-12-13
; PRIOR APPLICATION NUMBER: 09/669,445
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: PCT/US00/06642
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 60/126,018
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/139,638
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: 60/149,449
; PRIOR FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-013-477-20
```

```
Alignment Scores:
Pred. No.: 3.87e-18 Length: 40
Score: 32.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.16% Indels: 0
DB: Gaps: 0
```

US-10-030-271-3 (1-1883) x US-10-013-477-20 (1-40)

```
Qy 28 AAGATCCAAAGAGTGGTGGCGCGCTCGCCGAGGATCGAGCGGAGATCTGGC 87
Db 9 LysIleGlnAsnLysTrpLeuArgProSerProArgSerHisArgInProGlnSerIle 28
Qy 88 CGGGTCTGAGCTTGTCCGCTCTCCCTCCCGCGGA 123
Db 29 ArgValLeuSerLeuPheArgLeuProProGly 40
```

```
RESULT 11
US-09-733-167-6
; Sequence 6, Application US/09733167
; Patent No. US2002009909A1
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US2002009909A1 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C-DEED, which is a deletion mutant of human DEED comprising amino
; OTHER INFORMATION: acids 109-318 of the naturally occurring huma DEED.
US-09-733-167-6

Alignment Scores:
```

Pred. No.: 0.0553 Length: 210
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-733-167-6 (1-210)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 744
Db 84 ThrcysaspilleargleuargValArgAlaGluTrcys 96

RESULT 12
US-09-935-223-4
; Sequence 4, Application US/09935223
; Publication No. US20020086983A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, At
; FILE REFERENCE: TJU2499
; CURRENT APPLICATION NUMBER: US/09/935,223
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 09/723,450
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/276,993
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/859,167
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
; US-09-935-223-4

Alignment Scores:
Pred. No.: 0.0511 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-935-223-4 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 744
Db 192 ThrcysaspilleargleuargValArgAlaGluTrcys 204

RESULT 13
US-09-935-223-6
; Sequence 6, Application US/09935223
; Publication No. US20020086983A1
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, At
; FILE REFERENCE: TJU2499
; CURRENT APPLICATION NUMBER: US/09/935,223
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 09/723,450
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/276,993
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/859,167
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Novel Sequence
; US-09-935-223-6

Alignment Scores:
Pred. No.: 0.0511 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-935-223-6 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 744
Db 192 ThrcysaspilleargleuargValArgAlaGluTrcys 204

RESULT 14
US-09-733-167-1
; Sequence 1, Application US/09733167
; Patent No. US2002009009A1
; GENERAL INFORMATION:
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-733-167-1

Alignment Scores:
Pred. No.: 0.0511 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 9 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-733-167-1 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGAGCAGAGTACTGC 744
Db 192 ThrcysaspilleargleuargValArgAlaGluTrcys 204

RESULT 15
US-09-733-167-3
; Sequence 3, Application US/09733167
; Patent No. US2002009009A1
; GENERAL INFORMATION:
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. US2002009009A1 198 25 621.3

;; PRIOR FILING DATE: 1998-06-08
;; NUMBER OF SEQ ID NOS: 8
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 3
;; LENGTH: 318
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-733-167-3

Alignment Scores:

Pred. No.:	0 0511	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.10%	Indels:	0
DB:	9	Gaps:	0

US-10-030-271-3 (1-1883) x US-09-733-167-3 (1-318)

Qy 706 ACCTGTGACATCCGGCTCCGGTTCGAGCAGACTACTGC 744
Db 192 ThrcyAaprlaargleuargValargAlaGlutrcys 204

RESULT 16
US-10-437-963-107955
; Sequence 107955, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbasuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 107955
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_12256C.1.pep
US-10-437-963-107955

Alignment Scores:

Pred. No.:	3.26	Length:	107
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.79%	Indels:	0
DB:	16	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-437-963-107955 (1-107)

Qy 388 GTTCCCGCAGCCGCGAGTTGCTTCGCCGC 356
Db 17 ValAlaProAlaAlaAlaGlyCySerArgArg 27

RESULT 17
US-10-437-963-166242
; Sequence 166242, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

;; APPLICANT: Wu, Wei
;; APPLICANT: Boukharov, Andrey A.
;; APPLICANT: Barbasuk, Brad
;; APPLICANT: Li, Ping
;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
;; FILE REFERENCE: 38-21(53221)B
;; CURRENT APPLICATION NUMBER: US/10/437,963
;; CURRENT FILING DATE: 2003-05-14
;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 166242
;; LENGTH: 145
;; TYPE: PRT
;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_64970C.1.pep
US-10-437-963-166242

Alignment Scores:

Pred. No.:	3.08	Length:	145
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.77%	Indels:	0
DB:	16	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-437-963-166242 (1-145)

Qy 628 AGACGGCGCGGAGGGCGCCGACCCGCCACC 660
Db 51 ArgArgArgArgArgGlyAlaProAlaAlaPro 61

RESULT 18
US-10-001-254-38
; Sequence 38, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Krzysztof
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sung Hyung
; APPLICANT: Roth, Wilfried
; APPLICANT: Steiner-Bliwien, Frank
; TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
; FILE REFERENCE: P-1J 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-001-254-38

Alignment Scores:

Pred. No.:	22.1	Length:	146
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.61%	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-001-254-38 (1-146)

Qy 322 GAGCTCCTGCTGAGCTGAGCGCCGCCGG 351
Db 101 GluteuleuGlueuGlueuArgArgGly 110

RESULT 19
US-10-424-599-231701
; Sequence 231701, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 231701
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51249C.1.pep
US-10-424-599-231701

Alignment Scores:
Pred. No.: 186 Length: 66
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-424-599-231701 (1-66)

QY 344 CGCTCCAGCTCCAGAGAGCTCTAGG 318
DB 4 ArgSerSerSerArgSerSerArg 12

RESULT 20
US-10-767-701-47327
; Sequence 47327, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53235)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 47327
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C98_4.pep
US-10-767-701-47327

Alignment Scores:
Pred. No.: 175 Length: 90
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-767-701-47327 (1-90)

QY 628 AGACGGCGCGGAGAGGGGCCCGAGCC 654
DB 47 ArgArgArgArgArgGlyAlaProAla 55

RESULT 21
US-10-767-701-55826
; Sequence 55826, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53235)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 55826
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30933130.pep
US-10-767-701-55826

Alignment Scores:
Pred. No.: 173 Length: 94
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-767-701-55826 (1-94)

QY 301 GCCCGGGCCCGCAGCGGCTTAGAGCTC 327
DB 39 AlaArgAlaArgSerGlyLeuGluLeu 47

RESULT 22
US-10-424-599-255144
; Sequence 255144, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 255144
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72417C.1.pep
US-10-424-599-255144

Alignment Scores:
Pred. No.: 167 Length: 115
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-424-599-255144 (1-115)

QY 647 GCCCTCTCCGCGCGGCTTGACGACA 621
DB 45 AlaProLeuArgArgLeuAlaPro 53

RESULT 23
US-10-767-701-48342
; Sequence 48342, Application US/10767701
; Publication No. US2004012684A1
; GENERAL INFORMATION:
; APPLICANT: Kovallik, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 48342
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3481-027-P1-K1-B5.pep
US-10-767-701-48342

Alignment Scores:
Pred. No.: 167 Length: 116
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-767-701-48342 (1-116)

QY 344 CGCTCCAGCTCCAGCAGAGCTTAGG 318
DB 61 ArgSerSerSerArgSerSerArg 69

RESULT 24
US-10-437-963-133296
; Sequence 133296, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovallik, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133296
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35180C.1.pep
US-10-437-963-133296

Alignment Scores:
Pred. No.: 164 Length: 128
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-437-963-133296 (1-128)

QY 634 CGCGGAGAGGGCCCGCCAGCCGACCC 660
DB 2 ArgArgArgGlyAlaProAlaAlaPro 10

RESULT 25
US-10-437-963-142899
; Sequence 142899, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovallik, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142899
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43860C.1.pep
US-10-437-963-142899

Alignment Scores:
Pred. No.: 163 Length: 131
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-437-963-142899 (1-131)

QY 366 CAACCTGCGGCTGCGGCACTCCT 392
DB 48 GlnProAlaAlaAlaGlyAlaThrPro 56

RESULT 26
US-10-437-963-164498
; Sequence 164498, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovallik, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 164498
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_63391C.1.pep
US-10-437-963-164498

Alignment Scores:

Pred. No.:	162	Length:	133
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.45%	Indels:	0
DB:	16	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-437-963-164498 (1-133)

QY 792 GGGCTGGCGCGGACGTGAGCTGTT 818
DB 30 GlyAlaGlyAlaAlaAlaAlaGlyArgVal 38

RESULT 27

US-10-437-963-204745
; Sequence 204745, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 204745
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_99801C.1.pep
US-10-437-963-204745

Alignment Scores:

Pred. No.:	155	Length:	172
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	16	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-437-963-204745 (1-172)

QY 1166 CGTCTCCGAGAGAGTGCGCCGGA 1140
DB 25 ArgProIraArgIraArgIraArgIraArgIra 33

RESULT 28

US-09-833-245-369
; Sequence 369, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 369
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (64)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (113)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-833-245-369

Alignment Scores:

Pred. No.:	149	Length:	211
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	11	Gaps:	0

US-10-030-271-3 (1-1883) x US-09-833-245-369 (1-211)

QY 811 CCAGCTGCGCGCGGCGGCGGCGGCGG 785
DB 126 ProAlaAlaAlaProAlaProGlyAla 134

RESULT 29

US-10-437-963-124217
; Sequence 124217, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 124217
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_26979C.1.pep
US-10-437-963-124217

Alignment Scores:

Pred. No.:	148	Length:	215
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.45%	Indels:	0
DB:	16	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-437-963-124217 (1-215)

QY 1652 TCACCTTCCCGCTGCTGTCAGAGCT 1678
DB 29 SerProSerProLeuLeuSerArgAla 37

RESULT 30

US-10-425-114-56323
; Sequence 56323, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:

APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 56323
LENGTH: 216
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73131G05_FLI.pep
US-10-425-114-56323

Alignment Scores:
Pred. No.: 148 Length: 216
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x US-10-425-114-56323 (1-216)

QY 784 CGGCCCCAGCGCTGGCGCGCAGCTG 810
DB 164 ArgProGlnAlaLeuAlaArgGlnLeu 172

RESULT 31

US-10-437-963-197739
Sequence 197739, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 197739
LENGTH: 219
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_93469C.1.pep
US-10-437-963-197739

Alignment Scores:
Pred. No.: 148 Length: 219
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x US-10-437-963-197739 (1-219)

QY 786 GCGCCAGCGCTGGCGCGCAGCTGCA 812
|||||

DB 3 AlaProGlyAlaGlyAlaAlaGly 11

RESULT 32

US-10-437-963-175466
Sequence 175466, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 175466
LENGTH: 243
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_73308C.1.pep
US-10-437-963-175466

Alignment Scores:
Pred. No.: 145 Length: 243
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: Gaps: 0

US-10-030-271-3 (1-1883) x US-10-437-963-175466 (1-243)

QY 1064 TGGAGAGAGAGGCGCGCGCCGCA 1090
DB 74 TrpArgGlyGlyGlyGlyAlaArg 82
|||||

RESULT 33

US-10-211-028-31
Sequence 31, Application US/10211028
Publication No. US20050027113A1
GENERAL INFORMATION:
APPLICANT: CUBIST PHARMACEUTICALS, INC.
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO THE DAPTOMYCIN
FILE REFERENCE: CUB-12 PCT CIP
CURRENT APPLICATION NUMBER: US/10/211,028
CURRENT FILING DATE: 2002-07-31
PRIOR APPLICATION NUMBER: PCT/US02/24310
PRIOR FILING DATE: 2002-10-25
PRIOR APPLICATION NUMBER: PCT/US01/32354
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 60/310,385
PRIOR FILING DATE: 2001-08-06
PRIOR APPLICATION NUMBER: 60/379,866
PRIOR FILING DATE: 2002-05-10
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 262
TYPE: PRT
ORGANISM: Streptomyces roseosporus
US-10-211-028-31

Alignment Scores:
Pred. No.: 143 Length: 262
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 17 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-211-028-31 (1-262)

QY 930 CGCCCTGCTGCAGCGCCCTGGCGGCGT 956
DB 18 ArgProAlaAlaGlyProAlaGlyArg 26

RESULT 34
US-10-425-114-60272
Sequence 60272, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
NUMBER OF SEQ ID NOS: 2003-04-28
SEQ ID NO 60272
LENGTH: 298
TYPE: PRT
ORGANISM: Zea mays

FEATURE:
OTHER INFORMATION: Clone ID: LIB3600-053-C3_F11.pbp
US-10-425-114-60272

Alignment Scores:
Pred. No.: 139 Length: 298
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 15 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-425-114-60272 (1-298)

QY 622 GGTGCCAGACGGCGCGAGAGGGGCC 648
DB 74 GlyAlaHArgArgArgArgGlyAla 82

RESULT 35
US-10-437-963-104157
Sequence 104157, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 2003-05-14
SEQ ID NO 104157
LENGTH: 330
TYPE: PRT
ORGANISM: Oryza sativa

FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(330)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_101517C.1.pbp
US-10-437-963-104157

Alignment Scores:
Pred. No.: 137 Length: 330
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-437-963-104157 (1-330)

QY 792 GGCGCTGGCGCGAGCTGAGCTGTT 818
DB 177 GlyAlaGlyAlaAlaAlaGlyArgVal 185

RESULT 36
US-10-781-014-490
Sequence 490, Application US/10781014
Publication No. US20040180408A1
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Krogger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Habehauer, Gregor
TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
INVOLVED IN CARBON METABOLISM AND ENERGY
FILE REFERENCE: BGI-126CPN
CURRENT APPLICATION NUMBER: US/10/781,014
CURRENT FILING DATE: 2004-02-17
PRIOR APPLICATION NUMBER: US 09/602,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/143,208
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remaind Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 490
LENGTH: 336
TYPE: PRT
ORGANISM: Corynebacterium glutamicum

US-10-781-014-490

Alignment Scores:
Pred. No.: 136 Length: 336
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-781-014-490 (1-336)

Qy 967 GAGGCCCTGCGAGAGCTGTGGCCGG 993
58 GUAUAAUAUARGUAUAAUAGUAG 66

RESULT 37

US-10-285-976-1
Sequence 1, Application US/10285976
Publication No. US20030165500A1
GENERAL INFORMATION:

APPLICANT: Rhee, Chae-Seo
APPLICANT: Malini, Sen
APPLICANT: Wu, Christina
APPLICANT: Leon, Lorenzo M.
APPLICANT: Cort, MariPat
APPLICANT: Carson, Dennis A.
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Wnt and Frizzled Receptors as Targets for Immunotherapy
FILE REFERENCE: 023070-130320US
CURRENT APPLICATION NUMBER: US/10/285,976
CURRENT FILING DATE: 2002-11-01
PRIOR APPLICATION NUMBER: US 60/287,995
PRIOR FILING DATE: 2001-05-01
PRIOR APPLICATION NUMBER: WO PCT/US02/13802
PRIOR FILING DATE: 2002-05-01
NUMBER OF SEQ ID NOS: 232
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human Wnt-1
US-10-285-976-1

Alignment Scores:

Pred. No.:	134	Length:	370
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	14	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-285-976-1 (1-370)

Qy 366 GCTCTGCGCGCACTGCGCGCGCTC 340
18 AUAUAUAUAUAUAUAUAUAUAUAUA 26

RESULT 38

US-10-302-812-50
Sequence 50, Application US/10302812
Publication No. US20040087016A1
GENERAL INFORMATION:
APPLICANT: Keating et al.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CELL DIFFERENTIATION AND
FILE REFERENCE: HYDR-P02-004
CURRENT APPLICATION NUMBER: US/10/302,812
CURRENT FILING DATE: 2002-11-21
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.1
SEQ ID NO 50
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
US-10-302-812-50

Alignment Scores:

Pred. No.:	134	Length:	370
------------	-----	---------	-----

Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	15	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-302-812-50 (1-370)

Qy 366 GCTCTGCGCGCACTGCGCGCGCTC 340
18 AUAUAUAUAUAUAUAUAUAUAUAUA 26

RESULT 39

US-10-847-972-2
Sequence 2, Application US/10847972
Publication No. US20050049195A1
GENERAL INFORMATION:
APPLICANT: ZOU, YIMIN
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR NERVE REGENERATION
FILE REFERENCE: ARCD:395US
CURRENT APPLICATION NUMBER: US/10/847,972
CURRENT FILING DATE: 2004-05-17
PRIOR APPLICATION NUMBER: 60/470,913
PRIOR FILING DATE: 2003-05-15
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
US-10-847-972-2

Alignment Scores:

Pred. No.:	134	Length:	370
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.46%	Indels:	0
DB:	17	Gaps:	0

US-10-030-271-3 (1-1883) x US-10-847-972-2 (1-370)

Qy 366 GCTCTGCGCGCACTGCGCGCGCTC 340
18 AUAUAUAUAUAUAUAUAUAUAUAUA 26

RESULT 40

US-09-738-626-6644
Sequence 6644, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIRO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0

```
; SEQ ID NO 6644
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6644

Alignment Scores:
Pred. No.:      132      Length:      395
Score:          9.00      Matches:      9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    1.45%      Indels:      0
DB:             9          Gaps:        0

US-10-030-271-3 (1-1883) x US-09-738-626-6644 (1-395)

QY      967 GAGGCGCCGCGAGGCGTGTGGCGCG 993
DB      117 GUAAlaLeuArgGlnAlaValIGlYArg 125

RESULT 41
US-10-156-761-13365
; Sequence 13365, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13365
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13365

Alignment Scores:
Pred. No.:      132      Length:      398
Score:          9.00      Matches:      9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    1.45%      Indels:      0
DB:             14        Gaps:        0

US-10-030-271-3 (1-1883) x US-10-156-761-13365 (1-398)

QY      783 GCGGCGCCGCGCGCTGCGCGCGAGCT 809
DB      149 AlAlAlaProGlyAlaGlyAlaAlaAla 157

RESULT 42
US-10-425-114-38092
; Sequence 38092, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 38092
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3079-021-D5_FLI pep
US-10-425-114-38092

Alignment Scores:
Pred. No.:      131      Length:      410
Score:          9.00      Matches:      9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    1.45%      Indels:      0
DB:             15        Gaps:        0

US-10-030-271-3 (1-1883) x US-10-425-114-38092 (1-410)

QY      1417 ACACTGCCCGCTGATGATCTTCCCTG 1443
DB      17 TmLeuProLeuLeuArgSerSerLeu 25

RESULT 43
US-10-369-493-20883
; Sequence 20883, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20883
; LENGTH: 456
; TYPE: PRT
; ORGANISM: SYNECHOCOCCUS SP. WH 8102
US-10-369-493-20883

Alignment Scores:
Pred. No.:      128      Length:      456
Score:          9.00      Matches:      9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    1.46%      Indels:      0
DB:             15        Gaps:        0

US-10-030-271-3 (1-1883) x US-10-369-493-20883 (1-456)

QY      348 GCGGCGCTCCAGCTCCAGCAGAGAGCTC 322
DB      149 AlAlAlaLeuGlnLeuGlnGlnGlnLeu 157

RESULT 44
US-10-437-963-155657
; Sequence 155657, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 155657
LENGTH: 477
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_553C.1.dep
US-10-437-963-155657

Alignment Scores:
Pred. No.: 127 Length: 477
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 16 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-437-963-155657 (1-477)

QY 424 GCGGCGAGGCTGTCGCGGCGGCGGCA 398

DB 6 AAlaAlaAlaGlyArgGlyGlyProAla 14

RESULT 45

US-10-156-761-10964
Sequence 10964, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 10964
LENGTH: 478
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-10964

Alignment Scores:
Pred. No.: 127 Length: 478
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 14 Gaps: 0

US-10-030-271-3 (1-1883) x US-10-156-761-10964 (1-478)

QY 961 CTGACGAGGCGCTGCGAGAGGCTGCG 987

DB 25 LeuThnGluAlaLeuArgGluAlaVal 33

RESULT 46
US-08-459-455-36
Sequence 36, Application US/08459455
Publication No. US20030124105A1

GENERAL INFORMATION:
APPLICANT: Yuan, Junying
APPLICANT: Miura, Masayuki
TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSER: Sterne, Keesler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,455
FILING DATE: 2-JUN-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/368,704
FILING DATE: 4-JAN-1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/258,287
FILING DATE: 10-JUN-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/080,850
FILING DATE: 24-JUN-1993

ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.

REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609.3920003

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540
TELEX: 248636 SSX

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:

LENGTH: 503 amino acids
TYPE: amino acid
TOPOLOGY: both

US-08-459-455-36

Alignment Scores:
Pred. No.: 126 Length: 503
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 8 Gaps: 0

US-10-030-271-3 (1-1883) x US-08-459-455-36 (1-503)

QY 338 AGCTCAGAGAGGCTGAGGCGGCTG 312

DB 149 SerSerSerArgSerSerArgProLeu 157

RESULT 47

US-10-123-529-8
Sequence 8, Application US/10123529
Publication No. US20030105046A1

GENERAL INFORMATION:
APPLICANT: Yuan, Junying

Friedlander, Robert M.

TITLE OF INVENTION: Interleukin Converting Enzyme (ICE)
and Central Nervous System Damage
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/10/123,529
APPLICATION NUMBER DATA:
FILING DATE: 17-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/099,463A
FILING DATE: 18-Jun-1998
APPLICATION NUMBER: US 60/050,242
FILING DATE: 19-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Bugalsky, Lawrence B.
REGISTRATION NUMBER: 35,086
REFERENCE/DOCKET NUMBER: 0609.4420001/JAG/LBB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-123-529-8
Alignment Scores:
Pred. No.: 126 Length: 503
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 14 Gaps: 0
US-10-030-271-3 (1-1883) x US-10-123-529-8 (1-503)
QY 338 AGCTCCAGCAGGAGCTTAGCGCGCTG 312
DB 149 SerSerSerArgSerSerArgProleu 157
RESULT 48
US-09-888-243-5
Sequence 5, Application US/09888243
Patent No. US20020136714A1
GENERAL INFORMATION:
APPLICANT: Horvitz, H. Robert
APPLICANT: Yuan, Junying
APPLICANT: Shaham, Shai
TITLE OF INVENTION: Relatedness of Human Interleukin-1beta
TITLE OF INVENTION: Convertase Gene to a C. Elegans Cell Death Gene, Inhibitory
TITLE OF INVENTION: Portions of these Genes and Uses Therefor
FILE REFERENCE: 01997/211003
CURRENT APPLICATION NUMBER: US/09/888,243
CURRENT FILING DATE: 2001-06-22
PRIOR APPLICATION NUMBER: US 09/083,662
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: US 08/394,189
PRIOR FILING DATE: 1995-02-24

PRIOR APPLICATION NUMBER: US 08/282,211
PRIOR FILING DATE: 1994-07-11
PRIOR APPLICATION NUMBER: US 07/984,182
PRIOR FILING DATE: 1992-11-20
PRIOR APPLICATION NUMBER: US 07/897,788
PRIOR FILING DATE: 1992-06-12
NUMBER OF SEQ ID NOS: 30
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 505
TYPE: PRT
ORGANISM: Caenorhabditis briggsae
FEATURE:
NAME/KEY: VARIANT
LOCATION: 94, 95, 96, 120, 179, 318
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-888-243-5
Alignment Scores:
Pred. No.: 126 Length: 505
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 9 Gaps: 0
US-10-030-271-3 (1-1883) x US-09-888-243-5 (1-505)
QY 338 AGCTCCAGCAGGAGCTTAGCGCGCTG 312
DB 151 SerSerSerArgSerSerArgProleu 159
RESULT 49
US-10-425-114-62674
Sequence 62674, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 62674
LENGTH: 517
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3689-250-G2_FLI.pep
US-10-425-114-62674
Alignment Scores:
Pred. No.: 125 Length: 517
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 15 Gaps: 0
US-10-030-271-3 (1-1883) x US-10-425-114-62674 (1-517)
QY 347 CGCGGCTCCAGCTCCAGCAGGAGCTCT 321
DB 498 ArgArgSerSerSerSerArgSerSer 506
RESULT 50
US-10-425-114-42986

```
/ Sequence 42986, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313) B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 42986
/ LENGTH: 564
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: 700470960_FLI.pep
US-10-425-114-42986

Alignment Scores:
Pred. No.:      123      Length:      564
Score:          9.00      Matches:      9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      1.46%      Indels:      0
DB:              15      Gaps:      0

US-10-030-271-3 (1-1883) x US-10-425-114-42986 (1-564)
QY      347 CGCGCTCCAGCTCCAGCAGAGCTCT 321
Db      545 ArgArgSerSerSerArgSerSer 553

Search completed: March 22, 2005, 04:29:11
Job time : 529.739 secs
```

THIS PAGE LEFT BLANK

86 8 1.3 300 3 US-09-261-599B-6 Sequence 6, Appl1
87 8 1.3 300 3 US-09-261-599B-7 Sequence 7, Appl1
88 8 1.3 315 4 US-09-958-969-2 Sequence 2, Appl1
C 89 8 1.3 330 4 US-09-252-991A-28827 Sequence 28827, A
90 8 1.3 338 4 US-09-252-991A-31633 Sequence 31633, A
C 91 8 1.3 347 2 US-09-004-502-1 Sequence 1, Appl1
C 92 8 1.3 347 2 US-09-360-125-1 Sequence 1, Appl1
93 8 1.3 347 3 US-09-227-357-468 Sequence 468, App
C 94 8 1.3 348 4 US-09-252-991A-25605 Sequence 25605, A
95 8 1.3 354 4 US-09-949-016-11550 Sequence 11550, A
96 8 1.3 358 4 US-09-902-540-10833 Sequence 10833, A
97 8 1.3 361 3 US-09-261-599B-1 Sequence 1, Appl1
98 8 1.3 361 3 US-09-261-599B-4 Sequence 4, Appl1
99 8 1.3 361 4 US-09-456-455A-1 Sequence 1, Appl1
100 8 1.3 361 4 US-09-456-455A-4 Sequence 4, Appl1

ALIGNMENTS

RESULT 1

US-09-733-167A-6
; Sequence 6, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Krammer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: C-DEED, which is a deletion mutant of human DEED comprising amino
; OTHER INFORMATION: acids 109-318 of the naturally occurring human DEED.
US-09-733-167A-6

Alignment Scores:

Pred. No.: 0.00384 Length: 210
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-733-167A-6 (1-210)

QY 706 ACCTGTGACATCCGGCTCCGGGTTCCAGCAGACTACTGC 744
Db 84 ThrcysAspleArgLeuArgValArgAlaGluTyrCys 96

RESULT 2

US-08-859-167-4
; Sequence 4, Application US/08859167
; Patent No. 6037461
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s
; STREET: One Liberty Place, 46th floor

CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,167
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-859-167-4

Alignment Scores:

Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-08-859-167-4 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTCCAGCAGACTACTGC 744
Db 192 ThrcysAspleArgLeuArgValArgAlaGluTyrCys 204

RESULT 3

US-08-859-167-6
; Sequence 6, Application US/08859167
; Patent No. 6037461
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
; TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6037461r1s
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,167
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-

TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-859-167-6

Alignment Scores:
Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-08-859-167-6 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTCCAGCAGAGTACTGC 744
Db 192 ThrcYsaSpIleArGleuArGValArGAlaGluTrCys 204

RESULT 4

US-09-109-273-4
Sequence 4, Application US/09109273

PATENT No. 6063760
GENERAL INFORMATION:
APPLICANT: Alnemir, Emad S.
APPLICANT: Fernandez-Alnemir, Teresa
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
TITLE OF INVENTION: OF MAKING THE SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/109,273
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-109-273-4

Alignment Scores:
Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-109-273-4 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTCCAGCAGAGTACTGC 744
Db 192 ThrcYsaSpIleArGleuArGValArGAlaGluTrCys 204

RESULT 5

US-09-109-273-6
Sequence 6, Application US/09109273

PATENT No. 6063760
GENERAL INFORMATION:
APPLICANT: Alnemir, Emad S.
APPLICANT: Fernandez-Alnemir, Teresa
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
TITLE OF INVENTION: OF MAKING THE SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6063760r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/109,273
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-109-273-6

Alignment Scores:
Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-109-273-6 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTCCAGCAGAGTACTGC 744
Db 192 ThrcYsaSpIleArGleuArGValArGAlaGluTrCys 204

RESULT 6
US-09-276-993-4
Sequence 4, Application US/09276993

Patent No. 6207801
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
TITLE OF INVENTION: OF MAKING THE SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6207801r1s
STREET: One Liberty Place, 46th floor
City: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,993
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-276-993-4

Alignment Scores:
Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-276-993-4 (1-318)
Qy 706 ACCTGTGACATCGGCTCCGGGTTGAGCAGAGTACTGC 744
Db 192 ThrcyAspIleArgLeuArgValArgAlaGluIuTrCys 204

RESULT 7
US-09-276-993-6
Sequence 6, Application US/09276993
Patent No. 6207801
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: FADD-LIKE ANTI-APOPTOTIC MOLECULES, METHODS OF
TITLE OF INVENTION: USING THE SAME, AND COMPOSITIONS FOR AND METHODS
TITLE OF INVENTION: OF MAKING THE SAME
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 6207801r1s
STREET: One Liberty Place, 46th floor
City: Philadelphia
STATE: PA
COUNTRY: USA

ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: WINDOWS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276,993
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,167
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-276-993-6

Alignment Scores:
Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-276-993-6 (1-318)
Qy 706 ACCTGTGACATCGGCTCCGGGTTGAGCAGAGTACTGC 744
Db 192 ThrcyAspIleArgLeuArgValArgAlaGluIuTrCys 204

RESULT 8
US-09-723-450-4
Sequence 4, Application US/09723450
Patent No. 6576751
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, And
TITLE OF INVENTION: Compositions For And Methods Of Making The Same
FILE REFERENCE: TJU2445
CURRENT APPLICATION NUMBER: US/09/723,450
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 09/276,993
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 08/859,167
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 318
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: No. 6576751el Sequence
US-09-723-450-4

Alignment Scores:
Pred. No.: 0.00368 Length: 318
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.10% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-723-450-4 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGACGACGACTGTC 744
DB 192 ThrcyAsplleArGleuArGValArGAlaGluTrCyS 204

RESULT 9

US-09-723-450-6
; Sequence 6, Application US/09723450
; Patent No. 6576751
; GENERAL INFORMATION:
; APPLICANT: Alnemir, Emad S.
; TITLE OF INVENTION: Fadd-Like Anti-Apoptotic Molecules, Methods Of Using The Same, At
; TITLE OF INVENTION: Compositions For And Methods Of Making The Same
; FILE REFERENCE: RU2445
; CURRENT APPLICATION NUMBER: US/09/723,450
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/276,993
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/859,167
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Artificial Sequence
; NAME/KEY: misc feature
; OTHER INFORMATION: No. 6576751el Sequence
US-09-723-450-6

Alignment Scores:

Pred. No.:	0.00368	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.10%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-3 (1-1883) x US-09-723-450-6 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGACGACGACTGTC 744
DB 192 ThrcyAsplleArGleuArGValArGAlaGluTrCyS 204

RESULT 10

US-09-733-167A-1
; Sequence 1, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-733-167A-1

Alignment Scores:

Pred. No.:	0.00368	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.10%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-3 (1-1883) x US-09-733-167A-1 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGACGACGACTGTC 744
DB 192 ThrcyAsplleArGleuArGValArGAlaGluTrCyS 204

RESULT 11

US-09-733-167A-3
; Sequence 3, Application US/09733167A
; Patent No. 6696547
; GENERAL INFORMATION:
; APPLICANT: Peter, Marcus
; APPLICANT: Kramer, Peter
; TITLE OF INVENTION: Protein for Regulation of Apoptosis
; FILE REFERENCE: 4121-120
; CURRENT APPLICATION NUMBER: US/09/733,167A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: PCT/DE99/01712
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: German Patent Application No. 6696547 198 25 621.3
; PRIOR FILING DATE: 1998-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-733-167A-3

Alignment Scores:

Pred. No.:	0.00368	Length:	318
Score:	13.00	Matches:	13
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.10%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-3 (1-1883) x US-09-733-167A-3 (1-318)

QY 706 ACCTGTGACATCCGGCTCCGGGTTTCGACGACGACTGTC 744
DB 192 ThrcyAsplleArGleuArGValArGAlaGluTrCyS 204

RESULT 12

US-09-252-991A-30543
; Sequence 30543, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30543
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30543

Alignment Scores:

Pred. No.: 18.2 Length: 91
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-30543 (1-91)

Qy 628 AGACGGCGCGAGAGGCCCGCCAGCC 654
Db 5 ArgArgArgArgArgGlyAlaProAla 13

RESULT 13

US-09-252-991A-19469
; Sequence 19469, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19469
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19469

Alignment Scores:

Pred. No.: 17.3 Length: 148
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-19469 (1-148)

Qy 1166 CGTCCTCCGAGAGGTGCGCCGGA 1140
Db 104 ArgProProArgArgGTProGly 112

RESULT 14

US-09-252-991A-28027
; Sequence 28027, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28027
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28027

Alignment Scores: 16.2 Length: 273
Pred. No.: 16.2 Length: 273

Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-28027 (1-273)

Qy 599 GGAGTCGGGGCCGCCCACTGTGTG 625
Db 199 GlyValGlyAlaGlyProValVal 207

RESULT 15

US-09-902-540-12235
; Sequence 12235, Application US/09902540
; Patent No. 683447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12235
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12235

Alignment Scores:

Pred. No.: 16 Length: 305
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-902-540-12235 (1-305)

Qy 596 GGCGAGTCGGGGCCGCCCACTGTG 622
Db 236 GlyGlyValGlyAlaGlyProValVal 244

RESULT 16

US-09-252-991A-18524
; Sequence 18524, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18524
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18524

Alignment Scores: 15.9 Length: 331
Pred. No.: 15.9 Length: 331
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-18524 (1-331)

Qy 780 CCGCGGCGCCCGAGCGCTGCGCGCGCA 806
Db 124 ProAlaAlaProGlyAlaGlyAlaAla 132

RESULT 17
US-09-602-777A-230
; Sequence 230, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejun, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zeider, Oskar
; APPLICANT: Habershauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BG1-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6

; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 230
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-230

Alignment Scores:

Pred. No.: 15.8 Length: 336
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-602-777A-230 (1-336)

Qy 967 GAGCCCTCGAGAGAGCTGTGGCCG 993
Db 58 GAlaLeuArgGAlaAlaValGlyArg 66

RESULT 18
US-09-252-991A-18973
; Sequence 18973, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18973
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18973

Alignment Scores:

Pred. No.: 15.8 Length: 350
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-18973 (1-350)

Qy 634 CGGCGAGAGGCGCCGACCGCACCC 660
Db 119 ArgAlaGlyAlaProAlaAlaPro 127

RESULT 19

US-09-252-991A-18453
; Sequence 18453, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQUENCE ID NO: 18453
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18453

Alignment Scores:
Pred. No.: 15.7
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.45%
DB: 4
Length: 367
Matches: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-18453 (1-367)

QY 137 GGTCGACCCGCGCCGCTGCTGGAGG 163
Db 269 G1YArGrPrArGrPrArGrAlaG1YArG 277

RESULT 20
US-09-417-039-3
; Sequence 3, Application US/09417039A
; Patent No. 6485972
; GENERAL INFORMATION:
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Parr, Brian A.
; APPLICANT: Vaino, Seppo
; TITLE OF INVENTION: MYT SIGNALING IN REPRODUCTIVE ORGANS
; FILE REFERENCE: 00246/232001
; CURRENT APPLICATION NUMBER: US/09/417,039A
; CURRENT FILING DATE: 1999-10-12
; EARLIER APPLICATION NUMBER: US 60/109,355
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-417-039-3

Alignment Scores:
Pred. No.: 15.7
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.46%
DB: 4
Length: 370
Matches: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-09-417-039-3 (1-370)

QY 366 GCTCTGCGCGACGCTGCCGCGCGCTC 340
Db 18 AlaleuAlaAlaLeuProAlaAlaLeu 26

RESULT 21
US-09-252-991A-25124
; Sequence 25124, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQUENCE ID NO: 25124
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25124

Alignment Scores:
Pred. No.: 15.6
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.45%
DB: 4
Length: 386
Matches: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-25124 (1-386)

QY 622 GGTCGACGCGCGCGCGAGAGGCGCC 648
Db 171 G1YAlArGrArGrArGrArGrAla 179

RESULT 22
US-08-258-287B-36
; Sequence 36, Application US/08258287B
; Patent No. 6083735
; GENERAL INFORMATION:
; APPLICANT: Yuan, Junying
; APPLICANT: Miura, Masayuki
; TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,287B
; FILING DATE: 10-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/080,850
; FILING DATE: 24-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.3920001
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSX
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: both
US-08-258-287B-36

Alignment Scores:
Pred. No.: 15.2
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.46%
DB: 3
Length: 503
Matches: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0


```
US-10-030-271-3 (1-1883) x US-08-258-287B-36 (1-503)
Qy      338 AGCTCCAGCAGAGCTCTAGCCCGCTG 312
Db      149 SerSerSerArgSerSerArgProLeu 157

RESULT 23
US-08-368-704C-36
; Sequence 36, Application US/08368704C
; Patent No. 6087160
; GENERAL INFORMATION:
; APPLICANT: Yuan, Junying
; APPLICANT: Miura, Masayuki
; TITLE OF INVENTION: Programmed Cell Death Genes and Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/368,704C
; FILING DATE: 4-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,287
; FILING DATE: 10-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/080,850
; FILING DATE: 24-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0609.3920002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2540
; TELEFAX: (202) 371-2540
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; US-08-368-704C-36

Alignment Scores:
Pred. No.: 15.2      Length: 503
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.46%      Indels: 0
DB: 3      Gaps: 0

US-10-030-271-3 (1-1883) x US-08-368-704C-36 (1-503)
Qy      338 AGCTCCAGCAGAGCTCTAGCCCGCTG 312
Db      149 SerSerSerArgSerSerArgProLeu 157

RESULT 24
US-09-252-991A-23374
; Sequence 23374, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23374
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23374

Alignment Scores:
Pred. No.: 15.2      Length: 504
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.45%      Indels: 0
DB: 4      Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-23374 (1-504)
Qy      628 AGACGGCGGCGAGAGGGCCCGCAGCC 654
Db      16 ArgArgArgArgArgGlyAlaProAla 24

RESULT 25
US-08-394-189B-5
; Sequence 5, Application US/08394189B
; Patent No. 5962301
; GENERAL INFORMATION:
; APPLICANT: Horvitz, Robert
; APPLICANT: Yuan, Junying
; APPLICANT: Shahan, Shai
; TITLE OF INVENTION: RELATEDNESS OF HUMAN INTERLEUKIN-1
; TITLE OF INVENTION: BETA CONVERSEASE GENE TO A C. ELEGANS CELL DEATH
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Biring LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,189B
; FILING DATE: 24-FEB-2005
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/282,211
; FILING DATE: 12-JUL-1994
; APPLICATION NUMBER: 07/984,182
; FILING DATE: 20-NOV-1992
; APPLICATION NUMBER: 07/897,788
; FILING DATE: 12-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 01997/211001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
```

```
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-394-189B-5

Alignment Scores:
Pred. No.: 15.2 Length: 505
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x US-08-394-189B-5 (1-505)

QY 338 AGCTCCAGCAGAGCTTAGCGCGCTG 312
Db 151 SerSerSerArgSerSerArgProLeu 159

RESULT 26
PCT-US93-05701-20
; Sequence 20, Application PC/TUS9305701
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Cloning and Characterization of Cell Death Genes
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Massachusetts Institute of Technology
; STREET: 77 Massachusetts Avenue
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; COMPUTER:
; OPERATING SYSTEM:
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05701
; FILING DATE: 19930614
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: unsure
; LOCATION: at every Xaa
; PCT-US93-05701-20

Alignment Scores:
Pred. No.: 15.2 Length: 505
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 5 Gaps: 0

US-10-030-271-3 (1-1883) x PCT-US93-05701-20 (1-505)

QY 338 AGCTCCAGCAGAGCTTAGCGCGCTG 312
Db 151 SerSerSerArgSerSerArgProLeu 159

RESULT 27
PCT-US93-05705-5
```

```
; Sequence 5, Application PC/TUS9305705
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Inhibitors of Ced-3 and Related Proteins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Massachusetts Institute of Technology
; STREET: 77 Massachusetts Avenue
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05705
; FILING DATE: 19930714
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 505 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: unsure
; LOCATION: at every Xaa
; PCT-US93-05705-5

Alignment Scores:
Pred. No.: 15.2 Length: 505
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 5 Gaps: 0

US-10-030-271-3 (1-1883) x PCT-US93-05705-5 (1-505)

QY 338 AGCTCCAGCAGAGCTTAGCGCGCTG 312
Db 151 SerSerSerArgSerSerArgProLeu 159

RESULT 28
US-09-252-991A-28001
; Sequence 28001, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28001
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-28001

Alignment Scores:
Pred. No.: 14.7 Length: 657
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-252-991A-28001 (1-657)
```

OY 625 GCCAGCGCGCGAGAGGCCCA 651
| | | | | | | | | | | | | | | | | |
Db 88 AlargargargargargglyAlaPro 96

RESULT 29
US-08-630-916A-94
; Sequence 94, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K. M.
; APPLICANT: Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 866-8864/9741
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-916A-94

Alignment Scores:
Pred. No.: 178 Length: 15
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-08-630-916A-94 (1-15)

OY 402 GCCCGCCAGCAGCTGCTGCCGCA 425
| | | | | | | | | | | | | | | | | |
Db 7 GlyProProArgProAlaAlaAla 14

RESULT 30
US-08-602-999A-307
; Sequence 307, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.

APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; IDENTIFICATION AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 866-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 307:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-602-999A-307

Alignment Scores:
Pred. No.: 178 Length: 15
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-08-602-999A-307 (1-15)

OY 402 GCCCGCCAGCAGCTGCTGCCGCA 425
| | | | | | | | | | | | | | | | | |
Db 7 GlyProProArgProAlaAlaAla 14

RESULT 31
US-08-630-915A-150
; Sequence 150, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 150:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-150

Alignment Scores:
Pred. No.: 178 Length: 15
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 3 Gaps: 0

US-10-030-271-3 (1-1883) x US-08-630-915A-150 (1-15)
QY 402 GGCCGGCAGACCTGCTGCCGA 425
Db 7 GlyProPrArGrPrAlaAlaAla 14

RESULT 32
US-09-500-124-307
Sequence 307, Application US/09500124
Patent No. 6432920
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: OUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWIKES, Dana M.
APPLICANT: RIDER, James B.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,124
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/602,999
FILING DATE: 16-FEB-1996
ATTORNEY/AGENT INFORMATION:

NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 307:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-500-124-307

Alignment Scores:
Pred. No.: 178 Length: 15
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-500-124-307 (1-15)
QY 402 GGCCGGCAGACCTGCTGCCGA 425
Db 7 GlyProPrArGrPrAlaAlaAla 14

RESULT 33
US-09-879-957-150
Sequence 150, Application US/09879957
Patent No. 6709821
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6709821h
APPLICANT: KAY, Brian K.
APPLICANT: FOWIKES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/879,957
FILING DATE: 13-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,915
FILING DATE: 03-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Mastrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 150:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids

```

;
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 150:
US-09-879-957-150

Alignment Scores:
Pred. No.: 178 Length: 15
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-879-957-150 (1-15)

QY 402 GCGCCGCGCAGCAGCTGCTGCCGCA 425
Db 7 GlyProProArgProAlaAlaAla 14

RESULT 34
US-09-270-767-35900
; Sequence 35900, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35900
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-35900

Alignment Scores:
Pred. No.: 159 Length: 45
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-270-767-35900 (1-45)

QY 344 CGCTCCAGCTCCAGCAGAGCTCT 321
Db 10 ArgSerSerSerSerArgSerSer 17

RESULT 35
US-09-270-767-51117
; Sequence 51117, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 51117
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
```

```

US-09-270-767-51117

Alignment Scores:
Pred. No.: 159 Length: 45
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-270-767-51117 (1-45)

QY 344 CGCTCCAGCTCCAGCAGAGCTCT 321
Db 10 ArgSerSerSerSerArgSerSer 17

RESULT 36
US-08-726-306A-143
; Sequence 143, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Butsach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 143:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-726-306A-143

Alignment Scores:
Pred. No.: 156 Length: 52
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 2 Gaps: 0

US-10-030-271-3 (1-1883) x US-08-726-306A-143 (1-52)
```

QY 634 CGCGGAGAGGGGCCCCGACCGCA 657
DB 43 ArgArgArgGlyAlaProAlaAla 50

RESULT 37
US-09-061-026-16
; Sequence 16, Application US/09061026
; Patent No. 6077934
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, Richard
; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: Contryphan Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 755 Thirteenth Street N.W., Suite 701-E
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/061,026
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/068,737
; FILING DATE: 24-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2314-133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-061-026-16

Alignment Scores:
Pred. No.: 153 Length: 63
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-09-061-026-16 (1-63)

QY 1490 CTTAGGCCAAAGAAGTGTGCTG 1467
DB 51 LeuArgProIySylsCySValIeu 58

RESULT 38
US-09-466-138-16
; Sequence 16, Application US/09466138
; Patent No. 6153738
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, Richard
; APPLICANT: Olivera, Baldomero M.
; TITLE OF INVENTION: Contryphan Peptides
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 755 Thirteenth Street N.W., Suite 701-E

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Releasee #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/466,138
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/061,026
FILING DATE:
APPLICATION NUMBER: US 60/068,737
FILING DATE: 24-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 2314-133
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-783-6040
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-466-138-16

Alignment Scores:
Pred. No.: 153 Length: 63
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
Gaps: 0

US-10-030-271-3 (1-1883) x US-09-466-138-16 (1-63)

QY 1490 CTTAGGCCAAAGAAGTGTGCTG 1467
DB 51 LeuArgProIySylsCySValIeu 58

RESULT 39
US-09-489-039A-9033
; Sequence 9033, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Britton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIORITY FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9033
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9033

Alignment Scores:
Pred. No.: 153 Length: 65
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0

DB: 4 Gaps: 0
US-10-030-271-3 (1-1883) x US-09-489-039A-9033 (1-65)
QY 614 CCAGTGTGTGTCGACGAGCGGCG 637
Db 58 ProvalValValProAspGlyGly 65
RESULT 40
US-08-159-340A-6
; Sequence 6, Application US/06159340A
; Patent No. 5565352
; GENERAL INFORMATION:
; APPLICANT: Hochstrasser, Mark
; TITLE OF INVENTION: DEUBIQUITINATING ENZYME: COMPOSITIONS
; TITLE OF INVENTION: AND METHODS
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,340A
; FILING DATE: 24-NOV-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD.112/HYL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 67 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-159-340A-6
Alignment Scores:
Pred. No.: 152 Length: 67
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
Gaps: 0
US-10-030-271-3 (1-1883) x US-08-159-340A-6 (1-67)
QY 250 GAGCTCCTGGCCTTCTGCTGGAT 273
Db 33 GluLeuLeuAlaPheLeuLeuAsp 40
RESULT 41
US-09-489-039A-9947
; Sequence 9947, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9947
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9947
Alignment Scores:
Pred. No.: 150 Length: 76
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
Gaps: 0
US-10-030-271-3 (1-1883) x US-09-489-039A-9947 (1-76)
QY 268 CGCGGAGGCTTACCGGCGCG 311
Db 1 ArgAlaArgLeuSerProGlyPro 8
RESULT 42
US-09-902-540-14108
; Sequence 14108, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14108
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14108
Alignment Scores:
Pred. No.: 148 Length: 89
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.30% Indels: 0
Gaps: 0
US-10-030-271-3 (1-1883) x US-09-902-540-14108 (1-89)
QY 344 CGCTCAGCTCCAGCAGAGCTCT 321
Db 6 ArgSerSerSerArgSerSer 13
RESULT 43
US-09-489-039A-10032
; Sequence 10032, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747

;; PRIOR FILING DATE: 1999-01-29
;; NUMBER OF SEQ ID NOS: 14342
;; SEQ ID NO 10032
;; LENGTH: 98
;; TYPE: PRT
;; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10032

Alignment Scores:
Pred. No.: 146 Length: 98
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-489-039A-10032 (1-98)

QY 933 CCTGCTGAGGCGCTGGCGGCGT 956
DB 44 ProAlaGlyProAlaGlyArg 51

RESULT 44
US-09-513-999C-4550
; Sequence 4550, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4550
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -46...-1
; OTHER INFORMATION: score 7.8
; OTHER INFORMATION: seq LVLLGTRVPLSG/GP
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: -40
; OTHER INFORMATION: Xaa=asp or Glu
; NAME/KEY: UNSURE
; LOCATION: 22
; OTHER INFORMATION: Xaa=His or Asn
US-09-513-999C-4550

Alignment Scores:
Pred. No.: 146 Length: 101
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-513-999C-4550 (1-101)

QY 375 GCTGCTGGGCACTCTGCGCGT 398
DB 73 AlaAlaGlyAlaThrProAlaArg 80

RESULT 45

US-09-471-276-938
; Sequence 938, Application US/09471276
; Patent No. 6822072
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6822072
; FILE REFERENCE: GENSET.025CPI
; CURRENT APPLICATION NUMBER: US/09/471,276
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCT/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pm
; SEQ ID NO 938
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -46...-1
US-09-471-276-938

Alignment Scores:
Pred. No.: 146 Length: 101
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 4 Gaps: 0

US-10-030-271-3 (1-1883) x US-09-471-276-938 (1-101)

QY 375 GCTGCTGGGCACTCTGCGCGT 398
DB 73 AlaAlaGlyAlaThrProAlaArg 80

RESULT 46
US-09-949-016-7148
; Sequence 7148, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7148
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7148

Alignment Scores:
Pred. No.: 143 Length: 118
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.30% Indels: 0
DB: 4 Gaps: 0
US-10-030-271-3 (1-1883) x US-09-949-016-7148 (1-118)
QY 1090 TCGGGCGGGCCCCCTTCTCTCT 1067
DB 50 SerGlyAlaAlaProLeuProPro 57
RESULT 47
US-08-612-986-7
; Sequence 7, Application US/08612986
; Patent No. 5770384
; GENERAL INFORMATION:
; APPLICANT: Elliot J. Androphy
; APPLICANT: Dave E. Breiding
; TITLE OF INVENTION: E2 BINDING PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,986
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/361,806
; FILING DATE: 22 DEC 1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: NEP-004DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-612-986-7
Alignment Scores:
Pred. No.: 143 Length: 122
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 1 Gaps: 0
US-10-030-271-3 (1-1883) x US-08-612-986-7 (1-122)
QY 444 CCGGCCAGTGTCTCCAGACGCTA 467
DB 101 ProAlaSerValSerArgThrLeu 108
RESULT 48
US-08-361-806A-7
; Sequence 7, Application US/08361806A
; Patent No. 5792833
; GENERAL INFORMATION:
; APPLICANT: Elliot J. Androphy

APPLICANT: Dave E. Breiding
; TITLE OF INVENTION: E2 BINDING PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,806A
; FILING DATE: 22 DEC 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: NEP-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-361-806A-7
Alignment Scores:
Pred. No.: 143 Length: 122
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.29% Indels: 0
DB: 1 Gaps: 0
US-10-030-271-3 (1-1883) x US-08-361-806A-7 (1-122)
QY 444 CCGGCCAGTGTCTCCAGACGCTA 467
DB 101 ProAlaSerValSerArgThrLeu 108
RESULT 49
PCT-US95-16806A-7
; Sequence 7, Application PC/TUS9516806A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: E2 Binding Proteins
; NUMBER OF SEQUENCES: 21
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16806A
; FILING DATE: December 22, 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361,806
; FILING DATE: 22-DEC-1994
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

PCT-US95-16806A-7

Alignment Scores:

Pred. No.:	143	Length:	122
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	5	Gaps:	0

US-10-030-271-3 (1-1883) x PCT-US95-16806A-7 (1-122)

QY 444 CCGGCCAGTGTCTCCAGACCTA 467

DB 101 ProAlaSerValSerArgThrLeu 108

RESULT 50

US-09-252-991A-17914

/ Sequence 17914, Application US/09252991A

/ Patent No. 6551795

/ GENERAL INFORMATION:

/ APPLICANT: Marc J. Rubenfield et al.

/ TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS

/ FILE REFERENCE: 107196.136

/ CURRENT APPLICATION NUMBER: US/09/252,991A

/ CURRENT FILING DATE: 1999-02-18

/ PRIOR APPLICATION NUMBER: US 60/074,788

/ PRIOR FILING DATE: 1998-02-18

/ PRIOR APPLICATION NUMBER: US 60/094,190

/ PRIOR FILING DATE: 1998-07-27

/ NUMBER OF SEQ ID NOS: 33142

/ SEQ ID NO 17914

/ LENGTH: 133

/ TYPE: PRT

/ ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17914

Alignment Scores:

Pred. No.:	142	Length:	133
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.29%	Indels:	0
DB:	4	Gaps:	0

US-10-030-271-3 (1-1883) x US-09-252-991A-17914 (1-133)

QY 625 GCCAGACGCGCGCGAGAGGGGCC 648

DB 124 AlaArgArgArgArgGlyAla 131

Search completed: March 22, 2005, 03:23:12
Job time : 69.9542 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW nucleic - nucleic search, using sw model

Run on: March 23, 2005, 16:18:03 ; Search time 8475.52 Seconds
(without alignments)
10765.255 Million cell updates/sec

Title: US-10-030-271-3

Perfect score: 1883
Sequence: 1 aggcgcatacatagagagag.....ctcgaaagggcataggctg99 1883

Scoring table: OLIGO_NT
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_rtc:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1883	100.0	1883	6	BD093312 Apoptosis
2	1883	100.0	1883	9	AK075328 Homo sapi
3	1130	60.0	1230	6	AX322754 Sequence
4	1078	57.2	1924	6	AX431308 Sequence
5	973	51.7	1951	9	BC027930 Homo sapi
6	930	49.4	981	6	CQ719642 Sequence
7	909	48.3	909	9	AF443591 Homo sapi
8	909	48.3	909	6	CQ719642 Sequence
9	707	37.5	847	6	BD093311 Apoptosis
10	707	37.5	1659	9	AK130203 Homo sapi
11	707	37.5	2005	9	BC013372 Homo sapi
12	707	37.5	2012	9	AY125488 Homo sapi
13	707	37.5	95663	9	AC010247 Homo sapi
14	675	35.8	981	9	AF457575 Homo sapi
15	421	22.4	626	6	BD149917 Primer fo
16	421	22.4	626	6	AX869855 Sequence
17	368	19.5	1067	6	AR339362 Sequence
18	316	16.8	111084	6	AC006486 Homo sapi
19	268	14.2	603	6	AX136449 Sequence

20	268	14.2	603	6	BD123689	BD123689 Secretory
21	252	13.4	303	6	AX431298	AX431298 Sequence
22	237	12.6	441	6	CQ460291	CQ460291 Sequence
23	194	10.3	222	11	G20341	G20341 human STS A
24	187	9.9	1084	6	AX364851	AX364851 Sequence
25	187	9.9	1106	6	AX364852	AX364852 Sequence
26	165	8.8	381	6	BD154885	BD154885 Primer fo
27	165	8.8	381	6	AX874823	AX874823 Sequence
28	144	7.6	167108	2	AC068283	AC068283 Homo sapi
29	114	6.1	250	11	G31114	G31114 human STS S
30	62	3.3	226909	2	AC079490	AC079490 Mus muscu
31	56	3.0	993	10	AF457576	AF457576 Mus muscu
32	56	3.0	993	10	AF543541	AF543541 Mus muscu
33	56	3.0	1540	10	BC037043	BC037043 Mus muscu
34	56	3.0	190669	2	AC120393	AC120393 Mus muscu
35	41	2.2	226909	2	AC079490	AC079490 Mus muscu
36	37	2.0	1067	6	AR339362	AR339362 Sequence
37	37	2.0	1067	6	AR339362	AR339362 Sequence
38	24	1.3	197751	2	AC140075	AC140075 Rattus no
39	24	1.3	223642	2	AC108290	AC108290 Rattus no
40	23	1.2	1821	5	BC075935	BC075935 Danio rer
41	23	1.2	92932	9	AC091788	AC091788 Homo sapi
42	22	1.2	125523	5	BX248385	BX248385 Zebrafish
43	22	1.2	838	6	CQ411697	CQ411697 Sequence
44	22	1.2	1339	6	AX377832	AX377832 Sequence
45	22	1.2	2373	3	AF171096	AF171096 Strongylo
46	22	1.2	2718	6	CQ842142	CQ842142 Sequence
47	22	1.2	2718	6	AK125175	AK125175 Homo sapi
48	22	1.2	112782	2	AL355982	AL355982 Homo sapi
49	22	1.2	122619	2	AC143567	AC143567 Macaca mu
50	22	1.2	122976	2	AC143568	AC143568 Macaca mu
51	22	1.2	141272	9	AP001976	AP001976 Homo sapi
52	22	1.2	146371	2	AC151071	AC151071 Bos tauru
53	22	1.2	146691	2	AP002402	AP002402 Homo sapi
54	22	1.2	151235	9	AC090936	AC090936 Homo sapi
55	22	1.2	154937	2	AP002382	AP002382 Homo sapi
56	22	1.2	159577	2	AC080060	AC080060 Homo sapi
57	22	1.2	162088	2	AC069479	AC069479 Homo sapi
58	22	1.2	163028	9	AC005877	AC005877 c1b1 19 o
59	22	1.2	184026	9	AC0022176	AC0022176 Homo sapi
60	22	1.2	191824	9	AL161935	AL161935 Human DNA
61	22	1.2	195442	10	AL928974	AL928974 Mouse DNA
62	22	1.2	265792	2	AC134157	AC134157 Rattus no
63	22	1.2	265792	2	AC073684	AC073684 Mus muscu
64	21	1.1	21	6	AX431338	AX431338 Sequence
65	21	1.1	21	6	BD093317	BD093317 Apoptosis
66	21	1.1	23	6	BD093320	BD093320 Apoptosis
67	21	1.1	33	6	CQ406469	CQ406469 Sequence
68	21	1.1	1553	8	AK060228	AK060228 Oryza sat
69	21	1.1	1771	8	AK103120	AK103120 Oryza sat
70	21	1.1	2659	9	AK054721	AK054721 Homo sapi
71	21	1.1	2932	6	BD240835	BD240835 Isoforms
72	21	1.1	3234	6	BD240837	BD240837 Isoforms
73	21	1.1	3234	6	AR270869	AR270869 Sequence
74	21	1.1	3234	9	HSPCAR1	HSPCAR1 H. sapiens
75	21	1.1	40735	9	SB3176	SB3176 Homo sapien
76	21	1.1	62423	8	AL772208	AL772208 Human DNA
77	21	1.1	62423	8	AP006840_35	AP006840_35 Oryza sat
78	21	1.1	66135	1	AP006840_09	AP006840_09 Symlocbac
79	21	1.1	84792	2	AP005581	AP005581 Oryza sat
80	21	1.1	91642	2	AP005581	AP005581 Oryza sat
81	21	1.1	93919	2	AC002415	AC002415 Homo sapi
82	21	1.1	96095	4	BX640585	BX640585 Pig DNA
83	21	1.1	105521	8	AP005485	AP005485 Oryza sat
84	21	1.1	110000	1	AP006840_00	AP006840_00 Symlocbac
85	21	1.1	110000	1	AP006840_09	AP006840_09 Symlocbac
86	21	1.1	110000	1	AP006840_12	AP006840_12 Symlocbac
87	21	1.1	110000	1	AP006840_14	AP006840_14 Symlocbac
88	21	1.1	110000	1	AP006840_16	AP006840_16 Symlocbac
89	21	1.1	110000	1	AP006840_18	AP006840_18 Symlocbac
90	21	1.1	110000	1	AP006840_21	AP006840_21 Symlocbac
91	21	1.1	110000	1	AP006840_27	AP006840_27 Symlocbac
92	21	1.1	110000	1	AP006840_27	AP006840_27 Symlocbac

C 93	21	1.1 110000	1	AP006840_31	Continuation (32 o
C 94	21	1.1 110000	1	AP006840_33	Continuation (34 o
C 95	21	1.1 130054	8	AC133862	AC133862 Medicago
C 96	21	1.1 138141	8	AP005190	AP005190 Oryza sat
C 97	21	1.1 140624	8	AC124957	AC124957 Medicago
C 98	21	1.1 154537	8	AP005393	AP005393 Oryza sat
C 99	21	1.1 157026	2	AC120256	AC120256 Rattus no
C 100	21	1.1 164263	8	AP005515	AP005515 Oryza sat
ALIGNMENTS					
RESULT 1	BD093312	1883 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093312	Apoptosis Relating Factor.			
DEFINITION	BD093312.1	GI:22638900			
ACCESSION	BD093312.1	GI:22638900			
VERSION	WO 0104300-A/2.				
KEYWORDS	WO 0104300-A/2.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1883)				
JOURNAL	Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Miyoshi,S. and Sato,S. Apoptosis Relating Factor Patent: WO 0104300-A 2 18-JAN-2001; HELIX RESEARCH INSTITUTE, TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI, SUSUMU SATO				
COMMENT	OS Homo sapiens (human) PN WO 0104300-A/2 PD 18-JAN-2001 PF 06-JUL-2000 WO 2000JP004516 PR 08-JUL-1999 JP 99P 194179,18-OCT-1999 US 60/159586 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, YURI KAWAI, SOUSUKE MIYOSHI. P1 SUSUMU SATO				
FEATURES	PC C12N15/12, C07K14/47, C12N5/10, C12N1/21, C12N1/19, C12N1/15, C12P21/ PC 02', C07K16/18, C12P21/08, G01N33/53, G01N33/577 CC Key Location/Qualifiers FH Key Location/Qualifiers FT CDS Location/Qualifiers				
ORIGIN	1..1883 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"				
Query Match	100.0%; Score 1883; DB 6; Length 1883;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1883;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	AGGCGTAAATATAGAGAGGTGCAGAAAGATCCAAAACAAGTGCCTGCAGCCGCGCC	60		
Db	1	AGGCGTAAATATAGAGAGGTGCAGAAAGATCCAAAACAAGTGCCTGCAGCCGCGCC	60		
QY	61	AGGAGTCATGAGACGCCAGAACTGTGCGCGGTTCTGAGCTTTCCGCTCCCTCCCGG	120		
Db	61	AGGAGTCATGAGACGCCAGAACTGTGCGCGGTTCTGAGCTTTCCGCTCCCTCCCGG	120		
QY	121	GGAATGCGCTATCCCGGTCGACCCCGGCGCCGCTGCGGAGAGATGAGTGCCTGGAC	180		
Db	121	GGAATGCGCTATCCCGGTCGACCCCGGCGCCGCTGCGGAGAGATGAGTGCCTGGAC	180		
QY	181	TACTACGGATGCTGCTGCTTCAACCGTATGTTGAGGTGTGAGCGGCAACTGACCGAG	240		
Db	181	TACTACGGATGCTGCTGCTTCAACCGTATGTTGAGGTGTGAGCGGCAACTGACCGAG	240		
QY	241	TGCGAGCTGAGCTCTCTGCTTCTGCTGATGAGGCTCTGAGCGCCGCGGAGGCTTA	300		

Db	241	TGCGAGCTGAGCTCTCTGCTTCTGCTGATGAGGCTCTGAGCGCCGCGGAGGCTTA	300		
QY	301	GCCCCGGGCGCGAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCCGCGGAGGCTGAGC	360		
Db	301	GCCCCGGGCGCGAGCGGCTTAGAGCTCTGCTGAGCTGAGCGCCGCGGAGGCTGAGC	360		
QY	361	GAGAGCAACTGCGGCTGCTGCGGCAACTCTTCCGCTGCTGAGCCCGCAAGACTGCTG	420		
Db	361	GAGAGCAACTGCGGCTGCTGCGGCAACTCTTCCGCTGCTGAGCCCGCAAGACTGCTG	420		
QY	421	CCGCACTGCGCGCGAGCGGCGCGCGCGAGTCTCCAGAACGCTATGAGTATGCAAC	480		
Db	421	CCGCACTGCGCGCGAGCGGCGCGCGCGAGTCTCCAGAACGCTATGAGTATGCAAC	480		
QY	481	TCGAGCTTTCAAGAGAGCAGAGGCTAGCTCCGCTCGCGTCCGAGTCAAGAGTTC	540		
Db	481	TCGAGCTTTCAAGAGAGCAGAGGCTAGCTCCGCTCGCGTCCGAGTCAAGAGTTC	540		
QY	541	GCAAAATTCAGAGGCTCAGTGGAGACAGGCTCCCGCCAAACCAAGCGGACGCGCG	600		
Db	541	GCAAAATTCAGAGGCTCAGTGGAGACAGGCTCCCGCCAAACCAAGCGGACGCGCG	600		
QY	601	AGTCGGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660		
Db	601	AGTCGGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660		
QY	661	CAGCAGCACTGAGAGCGCGCGCAGACCTTCTCTGAGAGCAAGTGACTGTGACATCCG	720		
Db	661	CAGCAGCACTGAGAGCGCGCGCAGACCTTCTCTGAGAGCAAGTGACTGTGACATCCG	720		
QY	721	CTCCGGGCTTGAAGAGAGTATGCTGCGAGTGGGCCAGCTTGGAGAGCGGCGTGGATCC	780		
Db	721	CTCCGGGCTTGAAGAGAGTATGCTGCGAGTGGGCCAGCTTGGAGAGCGGCGTGGATCC	780		
QY	781	CGCGCGCGCGCGCGCTGCGCGCGCGCGCGCTGAGCGTGTGGGAGGCGCACCGGAGCTG	840		
Db	781	CGCGCGCGCGCGCGCTGCGCGCGCGCGCGCTGAGCGTGTGGGAGGCGCACCGGAGCTG	840		
QY	841	CGCTCAAGGAGACCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900		
Db	841	CGCTCAAGGAGACCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	900		
QY	901	GACGCTTCTGGGCGCACTTACCTGAGTGGCGCCCTGCTGAGCGCCCTGCGGCGCTTTC	960		
Db	901	GACGCTTCTGGGCGCACTTACCTGAGTGGCGCCCTGCTGAGCGCCCTGCGGCGCTTTC	960		
QY	961	CTGACTGAGGCGCTGCGAGGCTGTGGGCGGAGAGGCTGTGCGCTGCTGCTGCTGCTG	1020		
Db	961	CTGACTGAGGCGCTGCGAGGCTGTGGGCGGAGAGGCTGTGCGCTGCTGCTGCTGCTG	1020		
QY	1021	GATGAGCTGACTATGAGGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080		
Db	1021	GATGAGCTGACTATGAGGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1080		
QY	1081	CGCGCGCGCGCAGAGGCTCTGATTCAGAGACTGCGAGAGATTCACCTCCCAAGTCT	1140		
Db	1081	CGCGCGCGCGCAGAGGCTCTGATTCAGAGACTGCGAGAGATTCACCTCCCAAGTCT	1140		
QY	1141	CGCGGCGACCTTCTCGGAGAGAGAGCATCTTACCCCTTGAAGAGCCCTCCACAG	1200		
Db	1141	CGCGGCGACCTTCTCGGAGAGAGAGCATCTTACCCCTTGAAGAGCCCTCCACAG	1200		
QY	1201	ATGTGGGCTCTGAGGCTTAAACATTTCCAGCTGATTTCTTCCAGACTCTCTAC	1260		
Db	1201	ATGTGGGCTCTGAGGCTTAAACATTTCCAGCTGATTTCTTCCAGACTCTCTAC	1260		
QY	1261	CCGAGGTGCGCCCTTAAAGCTTCGAGAGCGGCGGCTGAGCTGTATCTCAGAGAGGAG	1320		
Db	1261	CCGAGGTGCGCCCTTAAAGCTTCGAGAGCGGCGGCTGAGCTGTATCTCAGAGAGGAG	1320		
QY	1321	GGCAGGCTACACCTCAACCAAGGCGCCCTGCAATTTGATCTTCTGAGTCTTGGGCTGT	1380		

Dh 1321 GGCACAGCTACACACTCACCAGGCCCCCTGACATTGATCTGATCTTGGGCTGT 1380
Qy 1381 CTGCACTGTACAGAGTGACACACACTGCTCATGTGTACACTGCGCCCTGTGAGATCTTCC 1440
Dh 1381 CTGCACTGTACAGAGTGACACACACTGCTCATGTGTACACTGCGCCCTGTGAGATCTTCC 1440
Qy 1441 CTGGGGCTCTGCCCTGTGGCTGCTTCCAGACACACTTCTTTGGCCCTTAAGGGCTTCTTC 1500
Dh 1441 CTGGGGCTCTGCCCTGTGGCTGCTTCCAGACACACTTCTTTGGCCCTTAAGGGCTTCTTC 1500
Qy 1501 TCAGACCTTATTTTGAACCAACCAACCTGGGCTTACCCACATATAGTGGGACCTGGA 1560
Dh 1501 TCAGACCTTATTTTGAACCAACCAACCTGGGCTTACCCACATATAGTGGGACCTGGA 1560
Qy 1561 GCTGGGGTGACATGGGGGCTGCTGCTTACCTTCCACATCTCCAGCAGCAGGGCCCT 1620
Dh 1561 GCTGGGGTGACATGGGGGCTGCTGCTTACCTTCCACATCTCCAGCAGCAGGGCCCT 1620
Qy 1621 GCCCAGCTTCAATTTTACAGACTGACTCTCTCTCACTTCCCTGCTGCTCAGAGCTGA 1680
Dh 1621 GCCCAGCTTCAATTTTACAGACTGACTCTCTCTCACTTCCCTGCTGCTCAGAGCTGA 1680
Qy 1681 ACATPAGCTTGCACTTGGATGTACCTGAGTGTACATGGAGTGTATAGCAGCATCA 1740
Dh 1681 ACATPAGCTTGCACTTGGATGTACCTGAGTGTACATGGAGTGTATAGCAGCATCA 1740
Qy 1741 TACCAAGGCTTACTGTTGCACTGGGGGCAAAACAGTAAACGCGCCTTCTTGAAG 1800
Dh 1741 TACCAAGGCTTACTGTTGCACTGGGGGCAAAACAGTAAACGCGCCTTCTTGAAG 1800
Qy 1801 GGAATGCAAAAGGCTTTGGGGGTATGTAAGAAAGCTTTTACAAATGATCAATTAAC 1860
Dh 1801 GGAATGCAAAAGGCTTTGGGGGTATGTAAGAAAGCTTTTACAAATGATCAATTAAC 1860
Qy 1861 GCCCTGGAAGGCGCATGTGGG 1883
Dh 1861 GCCCTGGAAGGCGCATGTGGG 1883

RESULT 2
AK075328 1883 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA PSEC0004 fls, clone NT2RM1000558, weakly similar
DEFINITION to Homo sapiens death effector domain-containing testicular
molecule mRNA.
AK075328
VERSION AK075328.1 GI:22761346
KEYWORDS oligo capping, fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Oca, T., Nishikawa, T., Suzuki, Y., Kawai, Hio, Y., Hayaashi, K.,
AUTHORS Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T.,
Nakamura, Y., Nagahari, K., Sugano, S. and Isogai, T.,
TITLE HRI human cDNA sequencing project
JOURNAL HRI human cDNA sequencing project
REFERENCE 2 (bases 1 to 1883)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hi.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT HRI human cDNA sequencing project, cDNA 5'- & 3'-end one pass
sequencing, clone selection and full insert sequencing: Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction: Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.

FEATURES

1. 1883
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="NT2RM1000558"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_id="NT2RM1"
/note="cloning vector: pUC19PL3-mRNA from uninduced NT2
neural precursor cells"
124..1104
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11551.1"
/db_xref="GI:22761347"
/translation="MALSGSTAPPCWEEBDECDLYGMLSHRMEFVVGQLTCEBEL
LAFIDBAPGAAGLARSGLELLEBRRCGCGESANLRLGLRLVLAHDLPL
ARKRRVPVPERYVGTSSSKRTESGSRROSSANSROGOMETSPTRORS
RGRPSGARRRRRRAAPQOQSEPARPSSEGVCTDILRVAYCEHSPALQCYA
SRPQALQADVFGQATVLSRSDYKSEVCDIKFSELSYLDLPKEDYLSGALLQALR
GVFLTALREAVGREAVRLVSVDEADYENGRRRLLMEEGGRPTPEAS"

ORIGIN

Query Match 100.0%; Score 1883; DB 9; Length 1883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCGTAAATATATAGAGAGGTGCGCAGAAAGATCCAAACAAAGTGTGCGCGCGCGCC 60
Dh 1 AGCGTAAATATATAGAGAGGTGCGCAGAAAGATCCAAACAAAGTGTGCGCGCGCGCC 60
Qy 61 AGGAGTCATGTGACGCGCAGAAATTTGGCGGGTCTGAGCTTGTTCGCGCTCTCCCGC 120
Dh 61 AGGAGTCATGTGACGCGCAGAAATTTGGCGGGTCTGAGCTTGTTCGCGCTCTCCCGC 120
Qy 61 AGGAGTCATGTGACGCGCAGAAATTTGGCGGGTCTGAGCTTGTTCGCGCTCTCCCGC 120
Dh 61 AGGAGTCATGTGACGCGCAGAAATTTGGCGGGTCTGAGCTTGTTCGCGCTCTCCCGC 120
Qy 121 GSAATGGCGCTATCCGGGTGACCCCGCGCCGCTGCTGGGAGGAGATGAGTCTGAGAC 180
Dh 121 GSAATGGCGCTATCCGGGTGACCCCGCGCCGCTGCTGGGAGGAGATGAGTCTGAGAC 180
Qy 181 TACTAGCGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGCGGCAATGACCGAG 240
Dh 181 TACTAGCGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGCGGCAATGACCGAG 240
Qy 241 TCGGAGCTGAGCTCTGCGCTTTTGTGTGATGAGGCTCTTGGCGCGCGGAGGCTTGA 300
Dh 241 TCGGAGCTGAGCTCTGCGCTTTTGTGTGATGAGGCTCTTGGCGCGCGGAGGCTTGA 300
Qy 301 GCCCGGGCCCGCAGCGGCTTGAAGCTCTGCTGAGAGTGAAGCGCGCGGCGAGTGGCGC 360
Dh 301 GCCCGGGCCCGCAGCGGCTTGAAGCTCTGCTGAGAGTGAAGCGCGCGGCGAGTGGCGC 360
Qy 361 GAGAGCAACTGCGGCTGTGGGCAACTCTCGGCTGTGGCCCGCAGACTGTGCTG 420
Dh 361 GAGAGCAACTGCGGCTGTGGGCAACTCTCGGCTGTGGCCCGCAGACTGTGCTG 420
Qy 421 CCGGACTGCGCGCAGCGGCGCGCGCGCAGTGTCTTCAGACGCTATAGCTATGCGACC 480
Dh 421 CCGGACTGCGCGCAGCGGCGCGCGCGCAGTGTCTTCAGACGCTATAGCTATGCGACC 480
Qy 481 TCCAGCTCTTCAAGGAGCAGAGGCTAGCTGCGGTGCGCGGTGGGAGTGAACAGTTCT 540
Dh 481 TCCAGCTCTTCAAGGAGCAGAGGCTAGCTGCGGTGCGCGGTGGGAGTGAACAGTTCT 540
Qy 541 GCAAAATTTTCAGAGGCTAGTGGAGACAGGCTCTCCCGCCCAACCAAGCGGCGCGCG 600
Dh 541 GCAAAATTTTCAGAGGCTAGTGGAGACAGGCTCTCCCGCCCAACCAAGCGGCGCGCG 600
Qy 601 AGTGGGGCGCGCGCAGTGTGTGTCAGACGCGCGCGGAGAGAGGCGCGCGCAGCC 660
Dh 601 AGTGGGGCGCGCGCAGTGTGTGTCAGACGCGCGCGGAGAGAGGCGCGCGCAGCC 660
Qy 661 CAGCAGAGTCAAGCGCGCAGACCTTCTCTGAAAGGCAAAATGACTGTGATCTTCGG 720
Dh 661 CAGCAGAGTCAAGCGCGCAGACCTTCTCTGAAAGGCAAAATGACTGTGATCTTCGG 720
Qy 721 CTCGGGTTTCAGACAGTACTGCGAGCATGTGGCCAGCTTGGAGCAGGCGTGGCATCC 780

|||||
Db 544 GCAAAATCTCAGCGAGGCTAGTGGAGAGCAGGCTCCCCCAACCAAGCCGACGCGCG 603
QY 601 AGTCGGGGCCGGCCAGTGTGTGTCCAGACGGGGGAGAGGGGCCCAAGCCGCAACC 660
Db 604 AGTCGGGGCCGGCCAGTGTGTGTCCAGACGGGGGAGAGGGGCCCAAGCCGCAACC 663
QY 661 CAGCAGAGTCAGAGCCCGCCAGACTTCTCTGAAGCAAAAGTACCTGTGATCCGG 720
Db 664 CAGCAGAGTCAGAGCCCGCCAGACTTCTCTGAAGCAAAAGTACCTGTGATCCGG 723
QY 721 CTCCGGGTTGAGAGAGTACTGCGAGATGGCCAGCTTGGAGCAGGGCGTGGCATCC 780
Db 724 CTCCGGGTTGAGAGAGTACTGCGAGATGGCCAGCTTGGAGCAGGGCGTGGCATCC 783
QY 781 CCGGGGGCCCAAGGCGTGGGCGGAGCTGAGCGTGTGGGCGAGGCCACCGCAGTCTG 840
Db 784 CCGGGGGCCCAAGGCGTGGGCGGAGCTGAGCGTGTGGGCGAGGCCACCGCAGTCTG 843
QY 841 CGCTCAAGGAGACTGGGCTGTGTGTGTGTGACATCAAGTCTGAGAGCTCTCTATCTG 900
Db 844 CGCTCAAGGAGACTGGGCTGTGTGTGTGTGTGACATCAAGTCTCTGAGAGCTCTCTATCTG 903
QY 901 GACGCTTCTGGGGCGACTACCTGATGCGCGCTCTGTCAGGCTCTGGGGGCGTGTTC 960
Db 904 GACGCTTCTGGGGCGACTACCTGATGCGCGCTCTGTCAGGCTCTGGGGGCGTGTTC 963
QY 961 CTGACTAGGCGCTTGGGAGAGGCTGTGGGCGGAGGCTGTTCCTGCTGTGTGATG 1020
Db 964 CTGACTAGGCGCTTGGGAGAGGCTGTGGGCGGAGGCTGTTCCTGCTGTGTGATG 1023
QY 1021 GATGAGGCTGACTATGAGGCTGCGCGGCGCTGTGTGTGATGAGAGAGAGGGGG 1080
Db 1024 GATGAGGCTGACTATGAGGCTGCGCGGCGCTGTGTGTGTGATGAGAGAGAGGGGG 1083
QY 1081 CGGCGCCCGACAGAGGCTCTCTGATCCAGAGCTGGCAGGATGTATCCACTCCAGTCT 1140
Db 1084 CGGCGCCCGACAGAGGCTCTCTGATCCAGAGCTGGCAGGATGTATCCACTCCAGTCT 1143
QY 1141 CCGGGGCACTTCTCTGGGAGAGCAACATCTTACCCCT 1181
Db 1144 CCGGGGCACTTCTCTGGGAGAGCAACATCTTACCCCT 1184

RESULT 4
AX431308
LOCUS AX431308 1924 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 17 from Patent WO240680.
ACCESSION AX431308
VERSION AX431308.1 GI:21656177
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS
1 Pawlowski, K., Fiorentino, L., Godzik, A., Lee, S. H., Reed, J. C.,
Roth, W. and Stenner-Lieman, F.
Novel death domain proteins
Patent: WO 0240680-A 17 23-MAY-2002;
BURNHAM INST (US)

FEATURES
source
1..1924
location/Qualifiers

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
91..>1044
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD37280.1"
/db_xref="GI:21656178"
/translation="WALSGSTPAPCWEDECLDYGMLSLHMFVVGQLTCELET
LAFLLDEAPAGGLARARSGLELLLELERRGQCDSESNRLTQLLAVLARHDLPHL

ARRRRPVSPEYRSYGTSSSKRTBGSCHRRRRSSSSANSQOQMETGSPHYRORRS
RGRPSGARRRRRGAAPAQOQSEPARPSSEGVTDIDLRVAFYCEHGPALQGV
SRPQALARQLDVGGATAVLRSDRSDSCDIXFSELSLADAFMPDYLSCALLQALR
GVFLTEALRAVREAVRLLVSVDEADYENGRRLLMBEQG"

Query Match 57.2%; Score 1078; DB 6; Length 1924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1128; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 CTTGCCAGAGATCATCGAGCCGAGATCTGGCCGGGTTCTGAGCTTGTCCGCTCC 112
Db 20 CGTCGCCAGAGATCATCGAGCCGAGATCTGGCCGGGTTCTGAGCTTGTCCGCTCC 79
QY 113 CTCCCCGGGAAATGGCGCTATCCGGGTCAGCCCGGCCCCGTGTGGAGAGAGATAGT 172
Db 80 CTCCCCGGGAAATGGCGCTATCCGGGTCAGCCCGGCCCCGTGTGGAGAGAGATAGT 139
QY 173 GCTGGACTACTACGGGATCTGTCTGCTTACCGTATGTTGAGGTGTGGCGGGCAAC 232
Db 140 GCTGGACTACTACGGGATCTGTCTGCTTACCGTATGTTGAGGTGTGGCGGGCAAC 199
QY 233 TGACCGAGTGCAGCTGGAAGCTCTTGCCCTTTTCTGTGATGAGGCTCTGGCGCCCG 292
Db 200 TGAACGAGTGCAGCTGGAAGCTCTTGCCCTTTTCTGTGATGAGGCTCTGGCGCCCG 259
QY 293 GAGGCTTAGCGCGGGCCCGAGCGGCTTAGAGCTCTGAGCTTGGAGCTGAGCGCGGGG 352
Db 260 GAGGCTTAGCGCGGGCCCGAGCGGCTTAGAGCTCTGAGCTTGGAGCTGAGCGCGGGG 319
QY 353 AGTCGCGAGAGCAACTGCGGCTGTGGGGCAACTCTGCGAGTGTGGCCGCCACG 412
Db 320 AGTCGCGAGAGCAACTGCGGCTGTGGGGCAACTCTGCGAGTGTGGCCGCCACG 379
QY 413 ACTTGCTGCGCACTTGGCGCGCGCAAGCGGCGCGCCCAAGTGTCTCCAGAACCTATAGCT 472
Db 380 ACTTGCTGCGCACTTGGCGCGCGCAAGCGGCGCGCCCAAGTGTCTCCAGAACCTATAGCT 439
QY 473 ATGGCACTTCAGTCTTCAAGAGGAGCAGAGGGTATGCTGCGGCGGCGGAGTCAA 532
Db 440 ATGGCACTTCAGTCTTCAAGAGGAGCAGAGGGTATGCTGCGGCGGCGGAGTCAA 499
QY 533 GCAATTCGCAAAATTCACAGCAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGC 592
Db 500 GCAATTCGCAAAATTCACAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGC 559
QY 593 AGCGGCGGAGTGGGGCGCGCCCAAGTGTGTGTCAGACGGCGCGGAGAGGGGCCACG 652
Db 560 AGCGGCGGAGTGGGGCGCGCCCAAGTGTGTGTCAGACGGCGCGGAGAGGGGCCACG 619
QY 653 CCGGACCCCGAGCAGAGTGCAGACCCCGCCAGACTTCTCTGAAGCAAAAGTGAACCTGTG 712
Db 620 CCGGACCCCGAGCAGAGTGCAGACCCCGCCAGACTTCTCTGAAGCAAAAGTGAACCTGTG 679
QY 713 ACAATCCGGCTCGGGTTGAGAGAGTACTGCAAGATGGGCCAGCTTGGAGCAGGGCG 772
Db 680 ACAATCCGGCTCGGGTTGAGAGAGTACTGCAAGATGGGCCAGCTTGGAGCAGGGCG 739
QY 773 TGGCATTCGGCGGGCCCGAGGCGTGGCGCGGCACTGGAAGTGTGGCGAGCCACCG 832
Db 740 TGGCATTCGGCGGGCCCGAGGCGTGGCGCGGCACTGGAAGTGTGGCGAGCCACCG 799
QY 833 CAGTGTGCGCTCAAGGAGCTTGGGGCTGTGTGTGTGTGACATCAAGTGTCTGAGAGCTT 892
Db 800 CAGTGTGCGCTCAAGGAGCTTGGGGCTGTGTGTGTGTGACATCAAGTGTCTGAGAGCTT 859
QY 893 CTTATCTGAGCGCTTCTGGGGCGACTACTGAGTGGCGGCTCTGTCAGAGCCCTGGCGG 952
Db 860 CTTATCTGAGCGCTTCTGGGGCGACTACTGAGTGGCGGCTCTGTCAGAGCCCTGGCGG 919
QY 953 GCGTGTCTGACTGAGGCGCTTGGAGAGGCTGTGGCGGGAGGCTGTTCGCTGTGG 1012
Db 920 GCGTGTCTGACTGAGGCGCTTGGAGAGGCTGTGGCGGGAGGCTGTTCGCTGTGG 979

QY 1013 TCAGTGTGATGAGGCTGATGAGGCTGCGCGCCGCTGTTGCTGATGAGAGAG 1072
 DB 980 TCAGTGTGATGAGGCTGATGAGGCTGCGCGCCGCTGTTGCTGATGAGAGAG 1039
 QY 1073 AAGCGGGGCGCGCCCGAGAGAGGCTCTGATTCAGAGACTGCGAGATTGATCCACCT 1132
 DB 1040 AAGCGGGGCGCGCCCGAGAGAGGCTCTGATTCAGAGACTGCGAGATTGATCCACCT 1099
 QY 1133 CCAAGTCTCCGGGCAACCTCTCTCGGAGAGAGACCATCTTACCCCT 1181
 DB 1100 CCAAGTCTCCGGGCAACCTCTCTCGGAGAGAGACCATCTTACCCCT 1148
 RESULT 5
 LOCUS BC027930
 DEFINITION Homo sapiens death effector domain containing 2, mRNA (cDNA clone
 MGC:35529 IMAGE:5199772), complete cds.
 ACCESSION BC027930
 VERSION BC027930.1 GI:20379818
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1951)
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Scheafer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
 Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S.,
 Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
 Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J.,
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
 Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Pahey J., Helton E., Ketterman M., Madan A.C., Shchepochko Y.,
 Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
 Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butlerfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E.,
 Scherch A., Schein J.E., Jones S.J. and Marra M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1951)
 Strausberg R.
 Direct Submission
 Submitted (08-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Ahther, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantipop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNLI at: <http://image.llnl.gov>
 Series: IRK Plate: 49 Row: b Column: 18
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 19923049.
 Location/Qualifiers
 1..1951

FEATURES

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:35529 IMAGE:5199772"
 /issue_type="Brain, adult, 6 pooled whole brains"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Vector: pCMV-Sport6"
 1..1951
 /gene="DED2"
 /note="synonym: FLAME-3"
 /db_xref="LOCUSID:162989"
 69..1049
 /gene="DED2"
 /codon_start=1
 /product="death effector domain-containing DNA binding
 protein 2"
 /protein_id="AAH27930.1"
 /db_xref="GI:20379819"
 /db_xref="LOCUSID:162989"
 /translation="MALSGTAPAEWEDECDIDYGMLSLHMFVVGQLTECELE
 LAFLEDEAGAGAGLARSGLLELLERRQCESSNRLGLQLRLVLRDLLEHL
 AKRRPVPERRYSYSSSTRTESSCRSSSNSSNOQOWETGSPPKRQRS
 RGRSPGARRRRRGAAPAOQSEPARPSEKVTCDILRYAAEYCEHGPLQGVVA
 SRPOLARQLDVPQATVLRSLSDISGVDCIKSELSYLAPFWDYVLSGLLQLRL
 GFLPTALRAEAVGREAVRLVSDVDADYAGRRRLLMEEBGRRTTAS"

CDS

gene

ORIGIN
 Query Match 51.7%; Score 973; DB 9; Length 1951;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1123; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 56 CGCCGAGAGTCATCGGAGCGCAAGTCTGCGCGGTTCTGAGCTTTCCTCCCTC 115
 DB 1 CGCCGAGAGTCATCGGAGCGCAAGTCTGCGCGGTTCTGAGCTTTCCTCCCTC 60
 116 CCGCGGAGTGGCGTATCGGCTGACCCCGGCGCGTGGGAGAGAGTATGAGCC 175
 DB 61 CCGCGGAGTGGCGTATCGGCTGACCCCGGCGCGTGGGAGAGAGTATGAGCC 120
 176 TGACTACTACGGGAGTGTGCTTCACTGATGTTGAGAGTGTGGCGGCACTGA 235
 DB 121 TGACTACTACGGGAGTGTGCTTCACTGATGTTGAGAGTGTGGCGGCACTGA 180
 236 CCGAGTGCAGAGTGTGAGTCTGCGCTTCTGCTGATGAGCTCTGCGCGCGGAG 295
 DB 181 CCGAGTGCAGAGTGTGAGTCTGCGCTTCTGCTGATGAGCTCTGCGCGCGGAG 240
 296 GCTTACCGCGCGCGCGCGAGCGGCTGAGCTCTGAGAGTGTGGCGCGGCGAGT 355
 DB 241 GCTTACCGCGCGCGCGCGAGCGGCTGAGAGTGTGGCGCGGCGGCGAGT 300
 356 GCGGAGAGAGCACTGCGGCTGCTGGGGCAACTCTGCGCGGCTGCGCGGCAAGACC 415
 DB 301 GCGGAGAGAGCACTGCGGCTGCTGGGGCAACTCTGCGCGGCTGCGCGGCAAGACC 360
 416 TGCTGCGGACCTGCGCGGCGCAAGCGGCGCGCACTGCTTCCAGAACCTATAGCTATG 475
 DB 361 TGCTGCGGACCTGCGCGGCGCAAGCGGCGCGCACTGCTTCCAGAACCTATAGCTATG 420
 476 GCACTTCAGACTTTCAAAGAGAGAGAGTGTGCTGCGCGGCTGCGCGGCAAGCA 535
 DB 421 GCACTTCAGACTTTCAAAGAGAGAGAGTGTGCTGCGCGGCTGCGCGGCAAGCA 480

QY 536 GTTCTCAATTCTCAGCAGAGGTCTAGTGGAGACAGGCTCCCCCAACCAAGCGGAGC 595
DB 481 GTTCTCAATTCTCAGCAGAGGTCTAGTGGAGACAGGCTCCCCCAACCAAGCGGAGC 540
QY 596 GCGGAGTCCGGGCGCGGCGGAGTGTGTGTCCAGACCGCGCGGAGGCGGCGGCGGCGG 655
DB 541 GCGGAGTCCGGGCGCGGCGGAGTGTGTGTCCAGACCGCGCGGAGGCGGCGGCGGCGG 600
QY 656 CACCCAGCAGCAGTCAAGCCCGCGAGCTTCTCTGAAGGAAAGTACCTGTGACA 715
DB 601 CACCCAGCAGCAGTCAAGCCCGCGAGCTTCTCTGAAGGAAAGTACCTGTGACA 660
QY 716 TCCGCTCCGGGTTCCAGCAGAGTACTGCGAGCATGGGCGGCGGCGGCGGCGGCGG 775
DB 661 TCCGCTCCGGGTTCCAGCAGAGTACTGCGAGCATGGGCGGCGGCGGCGGCGGCGG 720
QY 776 CATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 835
DB 721 CATCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
QY 836 TGTCTGCTCAAGGAGCCTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 895
DB 781 TGTCTGCTCAAGGAGCCTGGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 896 ATCTGACGCTTCTTGCGGCGGAGCTACTGAGTGGCGGCGGCGGCGGCGGCGGCGG 955
DB 841 ATCTGACGCTTCTTGCGGCGGAGCTACTGAGTGGCGGCGGCGGCGGCGGCGGCGG 900
QY 956 TGTCTGCTCAAGGAGCCTGGGCGGAGCTGTGGGCGGAGAGCTGTTCGCTGTGTGTCA 1015
DB 901 TGTCTGCTCAAGGAGCCTGGGCGGAGCTGTGGGCGGAGAGCTGTTCGCTGTGTGTCA 960
QY 1016 GT 1075
DB 961 GT 1020
QY 1076 GGGGCGGCGGCGGCGGAGAGGCTCTGTATCAAGAGCTGGAGAGATTGATCCCACTCCA 1135
DB 1021 GGGGCGGCGGCGGCGGAGAGGCTCTGTATCAAGAGCTGGAGAGATTGATCCCACTCCA 1080
QY 1136 AGTCTCCGGGCGGAGCTTCTCTGTGGAGAGCAGCATTCTTACCCCT 1181
DB 1081 AGTCTCCGGGCGGAGCTTCTCTGTGGAGAGCAGCATTCTTACCCCT 1126

RESULT 6
CQ719642 981 bp DNA 1linear PAT 03-FEB-2004
LOCUS CQ719642
DEFINITION Sequence 5576 from Patent WO02068579.
ACCESSION CQ719642
VERSION CQ719642.1 GI:42280499.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humaneurons or transcriptors, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 5576 06-SEP-2002;
FEATURES
source Location (NY) (US)
1. .981
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 49.4%; Score 930; DB 6; Length 981;
Best Local Similarity 99.9%; Pred. No. 0;

Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 124 ATGACGCTATCCGGGTCGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183
DB 1 ATGACGCTATCCGGGTCGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 184 TACGGAGT 243
DB 61 TACGGAGT 120
QY 244 GAGCTGAGAGTCTCTGGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 303
DB 121 GAGCTGAGAGTCTCTGGCTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 304 CGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 363
DB 181 CGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 364 AGCACTGCGGCTGTGGGCAACTCTGCGGCTGTGGGCGGCGGCGGCGGCGGCGG 423
DB 241 AGCACTGCGGCTGTGGGCAACTCTGCGGCTGTGGGCGGCGGCGGCGGCGGCGG 300
QY 424 CACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
DB 301 CACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 484 AGCTCTTCAAGAGGAGCAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
DB 361 AGCTCTTCAAGAGGAGCAGAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 544 AATTCTCAGCAGGAGT 603
DB 421 AATTCTCAGCAGGAGT 480
QY 604 CGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
DB 481 CGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 664 CAGCAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
DB 541 CAGCAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 724 CGGAGTTCAGCAGAGTACTGCGAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 783
DB 601 CGGAGTTCAGCAGAGTACTGCGAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 784 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843
DB 661 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 844 TCAAGGAGCCTGTGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 903
DB 721 TCAAGGAGCCTGTGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 904 GCTTCTGAGGAGCAGTACTGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 963
DB 781 GCTTCTGAGGAGCAGTACTGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 964 ACTGAGGCGGCTGTGAGAGGCTGTGGGCGGAGGAGGCTGTGTGTGTGTGTGTGTGT 1023
DB 841 ACTGAGGCGGCTGTGAGAGGCTGTGGGCGGAGGAGGCTGTGTGTGTGTGTGTGTGT 900
QY 1024 GAGGCTGAGCTATAGAGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1083
DB 901 GAGGCTGAGCTATAGAGTGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 1084 CGGCGGCGGAGGCGGCTCTCTGA 1104
DB 961 CGGCGGCGGAGGCGGCTCTCTGA 981

RESULT 7
AF443591

LOCUS	AF443591	981 bp	mRNA	linear	PRI 25-FEB-2002
DEFINITION	Homo sapiens death effector domain-containing DNA-binding protein 2 (DEDD) mRNA, complete cds.				
VERSION	AF443591				
KEYWORDS	AF443591.1	GI:17933266			
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 981)				
AUTHORS	Roth,W., Steiner-Lewen,F., Pawlowski,K., Godzik,A. and Reed,J.C.				
TITLE	Identification and characterization of DEDD2, a death effector domain-containing protein				
JOURNAL	J. Biol. Chem.	277 (9),	7501-7508	(2002)	
MEDLINE	21850646				
PUBMED	11741985				
REFERENCE	2 (bases 1 to 981)				
AUTHORS	Roth,W. and Reed,J.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-NOV-2001) Burnham Institute, 10901 N. Torrey Pines Rd, La Jolla, CA 92037, USA				
FEATURES	Location/Qualifiers				
source	1..981				

```

gene
CDS
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..981
/gene="DEDD2"
1..981
/gene="DEDD2"
/codon_start=1
/product="death effector domain-containing DNA-binding protein 2"
/protein_id="AA148220.1"
/db_xref="GI:17933267"
/translation="MALSGSTPAPCWEDECDLYGYGMLSLHMFVEVGGOLTECELEI
LAFLIDEPAGAGGLARASGLLEILLERFGQODENSLRLGLQLVLAHDHLLPHI
AKRRRPPSPRYSGTSSSSSKPTGSSCRRRRQSSSSNQQCGMEGSSPTKQRRR
RGRSGGRRRRRRGAPAPFOQSEPARSPSSGKTYCDRLKVRAPYCGHGFALGQVA
SRRFDALRLQLDVFGQATVAIVSRDLGSSVCDIFSELSYIDARWGDVLSQALLQALN
GVFLTEALREAVNGREAVRLIVSVSDAEYEAQRRLRLMEEEGGRPTKAS"

```

ORIGIN	Query Match	Similarity	49.4%	Score 930;	DB 9;	Length 981;
	Best Local	Similarity	99.9%	Pred. No. 0;		
	Matches	980;	Conservative	0;	Mismatches	1;
					Indels	0;
					Gaps	0;
QY	124	ATGGCGCTATCCGGGTGCAGCCCGGCCCGCTGCTGGAGAGAGATAGTGCCTGGACATAC				183
Db	1	ATGGGCGCTATCCGGGTGCAGCCCGGCCCGCTGCTGGAGAGAGATAGTGCCTGGACATAC				60
QY	184	TACGGGATGCTGTGCGCTTACCCGTAATGTTCCAGAGTGTGTGGCGCGGCAACTGACCGAATGTC				243
Db	61	TACGGGATGCTGTGCGCTTACCCGTAATGTTCCAGAGTGTGTGGCGCGGCAACTGACCGAATGTC				120
QY	244	GAGCTGAGAGCTCTGAGCCCTTTCTGTGCTGATGAGAGCTCCTGAGCGCGCCGAGAGGCTTAGCC				303
Db	121	GAGCTGAGAGCTCTGAGCCCTTTCTGTGCTGATGAGAGCTCCTGAGCGCGCCGAGAGGCTTAGCC				180
QY	304	CGGGCCGCGAGCGGCTTGAAGCTCTGCTGTGAGAGCTGAGACCGCGGCGCATGTGCGCGAG				363
Db	181	CGGGCCGCGAGCGGCTTGAAGCTCTGCTGTGAGAGCTGAGACCGCGGCGCATGTGCGCGAG				240
QY	364	AGCAACCTGCGCGCTGTGCTGGGGCAATCTCTGCGCGTGTGCGCCCGGCAAGACTGTGTCGCG				423
Db	241	AGCAACCTGCGCGCTGTGCTGGGGCAATCTCTGCGCGTGTGCGCCCGGCAAGACTGTGTCGCG				300
QY	424	CACCTGCGCGCCAAAGCGCGCGCGGCACTGTCTCAGAAAGCTAATAGCTAATGACACTCC				483
Db	301	CACCTGCGCGCCAAAGCGCGCGCGGCACTGTCTCAGAAAGCTAATAGCTAATGACACTCC				360
QY	484	AGCTTTCAAAAGAGACAGAGGGTAGTCCGCTGCGCCGTGGAGTCAAGACATTTCGCA				543

[illegible]

CC Key Location/Qualifiers
FH (1)..(909).
FT CDS Location/Qualifiers
FEATURES
source 1..909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 48.3%; Score 909; DB 6; Length 909;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 ATGGCGCTATCCGGGTGACCCCGCCGCGTGGGGAGGAGATGATGCTGGACTAC 183
DB 1 ATGGCGCTATCCGGGTGACCCCGCCGCGTGGGGAGGAGATGATGCTGGACTAC 60
QY 184 TACGGGATGCTGCTTCACCGTATGTTGAGGTGGGCGGCAACTGACGAGTGC 243
DB 61 TACGGGATGCTGCTTCACCGTATGTTGAGGTGGGCGGCAACTGACGAGTGC 120
QY 244 GAGCTGAGGCTCTGCGCTTTCTGCTGAGTGAAGGCTCTGCGCCGCGAGGCTTAACC 303
DB 121 GAGCTGAGGCTCTGCGCTTTCTGCTGAGTGAAGGCTCTGCGCCGCGAGGCTTAAGCC 180
QY 304 CGGGCCCGACGCGGCTAGAGCTCTGCTGAGGTGAGGCGCGCGGCGAGTGGCGGAG 363
DB 181 CGGGCCCGACGCGGCTAGAGCTCTGCTGAGGTGAGGCGCGCGGCGAGTGGCGGAG 240
QY 364 AGCAACTGCGGCTGCTGGGCACTCTGCGGCTGCGCCGCGCACTGCTGCGG 423
DB 241 AGCAACTGCGGCTGCTGGGCACTCTGCGGCTGCGCCGCGCACTGCTGCGG 300
QY 424 CACTGCGCGGCAAGCGGCGCCGCGCACTCTCAGAAAGCTAATGCTATGCACTCC 483
DB 301 CACTGCGCGGCAAGCGGCGCCGCGCACTCTCAGAAAGCTAATGCTATGCACTCC 360
QY 484 AGCTCTTCAAGAGACAGAGGAGTGCCTGCGGCTGCGGCACTAAGCACTTGCAG 543
DB 361 AGCTCTTCAAGAGACAGAGGAGTGCCTGCGGCTGCGGCACTAAGCACTTGCAG 420
QY 544 AATTCTCAGAGGCTCAGTGGGAGACAGGCTCTCCCGCAACCAAGCGGCGGAGT 603
DB 421 AATTCTCAGAGGCTCAGTGGGAGACAGGCTCTCCCGCAACCAAGCGGCGGAGT 480
QY 604 CGGGGCGGCGCACTGCTGCTGCGCAAGCGGCGGAGAGGCGCCGCACTGCGG 663
DB 481 CGGGGCGGCGCACTGCTGCTGCGCAAGCGGCGGAGAGGCGCCGCACTGCGG 540
QY 664 CAGCAGTCAGAGCGCGCAGACCTTCTCTGAAGGCAAGTGAACCTGTCATCCGCTC 723
DB 541 CAGCAGTCAGAGCGCGCAGACCTTCTCTGAAGGCAAGTGAACCTGTCATCCGCTC 600
QY 724 CGGGTTTCAAGAGTACTGCGAGCATGAGGCGCAGCTTTGAGAGAGGCGTGGCATCCCG 783
DB 601 CGGGTTTCAAGAGTACTGCGAGCATGAGGCGCAGCTTTGAGAGAGGCGTGGCATCCCG 660
QY 784 CGGCCCCAGGCGCTGGCGCGGCGAGCTGAGCTGTGTTGGGCGACCGAGTGTGCGC 843
DB 661 CGGCCCCAGGCGCTGGCGCGGCGAGCTGAGCTGTGTTGGGCGACCGAGTGTGCGC 720
QY 844 TCAAGGAGACCTGGGCTCTGTGTTTGAATCAAGTTCAGAGCTCTCTATCTGGAG 903
DB 721 TCAAGGAGACCTGGGCTCTGTGTTTGAATCAAGTTCAGAGCTCTCTATCTGGAG 780
QY 904 GCGCTTCTGGGCGCACTAAGTGTGCGCCCTGCTGAGGCGCTGCGGCGGCTGTTCTCTG 963
DB 781 GCGCTTCTGGGCGCACTAAGTGTGCGCCCTGCTGAGGCGCTGCGGCGGCTGTTCTCTG 840
QY 964 ACTGAGGCGCTTGGAGAGGCTGTGGGCGGAGAGGCTTTGCTGCTGCTGCTGAGT 1023
DB 841 ACTGAGGCGCTTGGAGAGGCTGTGGGCGGAGAGGCTTTGCTGCTGCTGCTGAGT 900

QY 1024 GAGGCTGAC 1032
DB 901 GAGGCTGAC 909

RESULT 9
AR379637
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source 1..847
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 37.5%; Score 707; DB 6; Length 847;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1177 CCCCTTGACAGCCCTCCCAAGATGTGGGCTCTGAGGCTAAACCATTTCCAGCTGAG 1236
DB 141 CCCCTTGACAGCCCTCCCAAGATGTGGGCTCTGAGGCTAAACCATTTCCAGCTGAG 200
QY 1237 TTTCTTCCCAAGCTCTCTTACCCCAAGTGTGCCCCCTTTAGCTCCGAGGCGGGG 1296
DB 201 TTTCTTCCCAAGCTCTCTTACCCCAAGTGTGCCCCCTTTAGCTCCGAGGCGGGG 260
QY 1297 TGGGCTGTATCTCAGAGAGGAGGCGACAGCTACACCTCAACCAAGGCCCCCTG 1356
DB 261 TGGGCTGTATCTCAGAGAGGAGGCGACAGCTACACCTCAACCAAGGCCCCCTG 320
QY 1357 ATTGTATCTGTATCTTGGGCTGTGCACTGTACAGGTCACACCTGCTCATGCTC 1416
DB 321 ATTGTATCTGTATCTTGGGCTGTGCACTGTACAGGTCACACCTGCTCATGCTC 380
QY 1417 ACACTGCCCCCTGCTGAGATCTTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 1476
DB 381 ACACTGCCCCCTGCTGAGATCTTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 440
QY 1477 TTTCTTGGCTTAAGGCTTCTCTCTCAGGACCTCTAATTGACCAACCACTGGGCT 1536
DB 441 TTTCTTGGCTTAAGGCTTCTCTCTCAGGACCTCTAATTGACCAACCACTGGGCT 500
QY 1537 TCAGGCACATCAGTGGGCACTGAGGCTGGGGTSCAATGGGGCTGTGCTGCTGCTG 1596
DB 501 TCAGGCACATCAGTGGGCACTGAGGCTGGGGTSCAATGGGGCTGTGCTGCTGCTG 560
QY 1597 ACATCTTCAGCCAGCAGGCGCCCTGCGCAGCTTCAATTATACAGACCTGACTCCACC 1656
DB 561 ACATCTTCAGCCAGCAGGCGCCCTGCGCAGCTTCAATTATACAGACCTGACTCCACC 620
QY 1657 TTTCCCCCTGCTCTCAGAGCTGAATAGACTTGGACCTTGGATGTCACTGAGTGTCA 1716
DB 621 TTTCCCCCTGCTCTCAGAGCTGAATAGACTTGGACCTTGGATGTCACTGAGTGTCA 680
QY 1717 CATGGAGTGTATGGCAGATCATCAAGGCTTCTGTTGCAATGGGGCCAAACCA 1776
DB 681 CATGGAGTGTATGGCAGATCATCAAGGCTTCTGTTGCAATGGGGCCAAACCA 740
QY 1777 GTAACAGCCACTTTCTTGGAAAGGAAATGCAAGGCTTTGGGGGTGATGAAAGACCT 1836

Db 741 GTAAACACCCCTTTCTTGAAAGGAGATCAAGGCTTTGGGGGTGATGGAAGAAAGACT 800

Qy 1837 TTTCACAAATGATACCAATTAACCTGCTTGAAAGGAGCATAGGTGG 1883

Db 801 TTTCACAAATGATACCAATTAACCTGCTTGAAAGGAGCATAGGTGG 847

RESULT 10
AK130203 1659 bp mRNA linear PRI 10-SEP-2003
LOCUS Homo sapiens cDNA FLJ26693 fis, clone MFG07953.
DEFINITION AK130203
ACCESSION AK130203.1 GI:34526957
VERSION oligo capping, fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS 1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H., Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A., Kawakami, B., Nagai, K., Isogai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1659)
AUTHORS Sugano, S. and Suzuki, Y.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: fldcna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan, cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.

FEATURES
source
1..1659
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MFG07953"
/cell_type="macrophage"
/clone_1lb="WPG"
/note="cloning vector: pME18SFL3-differentiated by oxidant"

ORIGIN
Query Match 37.5%; Score 707; DB 9; Length 1659;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1177 CCCCTTGACAGCCCTCCACAGAGATGTGGCTCTGAGGCTTAACATTTCCAGCTGAG 1236

Db 952 CCCCTTGACAGCCCTCCACAGAGATGTGGCTCTGAGGCTTAACATTTCCAGCTGAG 1011

Qy 1237 TTTCTTTCCAGACTCTCTCTACCCAGGTGTGCCCCCTTAAGCTTCGGAGCGGGGCG 1296

Db 1012 TTTCTTTCCAGACTCTCTCTACCCAGGTGTGCCCCCTTAAGCTTCGGAGCGGGGCG 1071

Qy 1297 TGGGCGCTGATCTTCAGAAAGGAGGGGACAGCTACACACTCAACAAAGGCCCTTGAC 1356

Db 1072 TGGGCGCTGATCTTCAGAAAGGAGGGGACAGCTACACACTCAACAAAGGCCCTTGAC 1131

Qy 1357 ATTGTATCTGTGATCTTTGGGCTGTCTGCACTGTCAAGGTGACACACTGCTCATGCTC 1416

Db 1132 ATTGTATCTGTGATCTTTGGGCTGTCTGCACTGTCAAGGTGACACACTGCTCATGCTC 1191

Qy 1417 ACACTGCCCCCTGTGTGATCTTCCCTGGGCTCTGCGCCCTGCTGTTCCACACACAC 1476

Db 1192 ACACTGCCCCCTGTGTGATCTTCCCTGGGCTCTGCGCCCTGCTGTTCCACACACAC 1251

Qy 1477 TTCTTTGGCTTAAGGGCTCTCTCTGAGACCTCTTAATTGACCAACAACACTGGGCT 1536

Db 1252 TTCTTTGGCTTAAGGGCTCTCTCTGAGACCTCTTAATTGACCAACAACACTGGGCT 1311

Qy 1537 TCAGCCACATCAGTGGGACCTGAGAGCTGGGGTGCACATGGGGCTGCTGACCTTGCCAC 1596

Db 1312 TCAGCCACATCAGTGGGACCTGAGAGCTGGGGTGCACATGGGGCTGCTGACCTTGCCAC 1371

Qy 1597 ACATCTTCAGCCACGAGGCGCCCTGCGGCTTCAATTTACAGACCTGACTCTCCACAC 1656

Db 1372 ACATCTTCAGCCACGAGGCGCCCTGCGGCTTCAATTTACAGACCTGACTCTCCACAC 1431

Qy 1657 TTCCCCCTGCTGTCCAGAGCTGAACATAGACTTGACCTTGATGTCACCTGGAGTGTCA 1716

Db 1432 TTCCCCCTGCTGTCCAGAGCTGAACATAGACTTGACCTTGATGTCACCTGGAGTGTCA 1491

Qy 1717 CATGGAGGTGTTATGAGCAGCATCATCAAGGCTTACTGTTGACATGGGGCCAAACCA 1776

Db 1492 CATGGAGGTGTTATGAGCAGCATCATCAAGGCTTACTGTTGACATGGGGCCAAACCA 1551

Qy 1777 GTAAACAGCCACTCTTTTGAAAGGAGATGCAGAAAGCTTTGGGGGTGATGAAAGACCT 1836

Db 1552 GTAAACAGCCACTCTTTTGAAAGGAGATGCAGAAAGCTTTGGGGGTGATGAAAGACCT 1611

Qy 1837 TTTCACAAATGATACCAATTAACCTGCTTGAAAGGAGCATAGGTGG 1883

Db 1612 TTTCACAAATGATACCAATTAACCTGCTTGAAAGGAGCATAGGTGG 1658

RESULT 11
BC013372 2005 bp mRNA linear PRI 24-FEB-2004
LOCUS BC013372
DEFINITION Homo sapiens death effector domain containing 2, mRNA (cDNA clone MGC:16414 IMAGE:3941907), complete cds.
ACCESSION BC013372
VERSION BC013372.2 GI:33872465
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2005)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schertz, T.E., Brownstein, M.J., Uedln, T.B., Toshyluk, S., Garincci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, D.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smallu, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2005)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK
COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15426522.
Contact: MGC help desk
Email: cgapsd-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
Contact: amadad@systemsbiology.org
Anup Madan, Jeessica Fahey, Erin Heltton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>
Series: IRAL Plate: 25 Row: F Column: 8.
Location/Qualifiers

FEATURES
source

1..2005
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:16414 IMAGE:3941907"
/issue_type="Lung, small cell carcinoma"
/clone_id="NIH MGC 7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1..2005
/gene="DDEDD2"
/note="synonym: FLAME-3"
/db_xref="LocusID:162989"
75..632
/gene="DDEDD2"
/codon_start=1
/product="DDEDD2 protein"
/protein_id="AAH13372.2"
/db_xref="GI:33872466"
/db_xref="LocusID:162989"
/translation="MALSGSTPAPCWEDEBCLDYGMLSHRFVVGQLTCELEL
LAFLDARPAAGLARASGLLELELRGGQDSESNLILGQLVLRHDLPLH
ARRRRPVSPEPERSYGTSSSKRTGSCRRRRSSSSANSQGSPPTKRRRSRGRPS
GGARRRRRGAAPAPQOQSEPAQTFL"

CDS

gene

ORIGIN

Query Match 37.5%; Score 707; DB 9; Length 2005;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1177 CCCCTTGACAGCCCTCCCAAGAGATGTGGGCTGTGAGGCTTAAACATTTCCAGCTGAG 1236
1174 CCCCTTGACAGCCCTCCCAAGAGATGTGGGCTGTGAGGCTTAAACATTTCCAGCTGAG 1233
1237 TTTCCTTCCGAGACTCCTCTTACCCCGAGGTGTGCCCCCTTAGCTCCGAGGCGGGGC 1296
1234 TTTCCTTCCGAGACTCCTCTTACCCCGAGGTGTGCCCCCTTAGCTCCGAGGCGGGGC 1293
1297 TGGGCGCTGATATTCAGAAAGGAGGAGGACAGCTACACACTCAACAAAGGCCCTTCAC 1356
1294 TGGGCGCTGATATTCAGAAAGGAGGAGGACAGCTACACACTCAACAAAGGCCCTTCAC 1353
1357 ATTGATATCTGATCTTGGGCTGTCTGACATGTGACAGGTGACACACTGCTCATGCTC 1416
1354 ATTGATATCTGATCTTGGGCTGTCTGACATGTGACAGGTGACACACTGCTCATGCTC 1413
1417 ACACTGCGCTGTGAGATCTTCCCTGAGGCTCTGCCCCCTGCTTCCCAAGACACAC 1476
1414 ACACTGCGCTGTGAGATCTTCCCTGAGGCTCTGCCCCCTGCTTCCCAAGACACAC 1473
1477 TTCTTTGGCCTAAGGGGCTTCTCTCTCAGACCTCTAATTTGAACCAACCAACTGGGCT 1536
1474 TTCTTTGGCCTAAGGGGCTTCTCTCTCAGACCTCTAATTTGAACCAACCAACTGGGCT 1533

1537 TCAGCCACATCAGTGGGACCTGAGGCTGGGTCACATGAGGCGCTGCTACCTTGGCCAC 1596
1534 TCAGCCACATCAGTGGGACCTGAGGCTGGGTCACATGAGGCGCTGCTACCTTGGCCAC 1593
1597 ACATCTTCAGCCAGCCAGGCGCTTGGCCAGCTTCAATTTACAGACTGATCTTCTCACC 1656
1594 ACATCTTCAGCCAGCCAGGCGCTTGGCCAGCTTCAATTTACAGACTGATCTTCTCACC 1653
1657 TTTCCCGCTGCTTCAGAGCTGAATGACATGTCACCTTGAGTGCATGAGAGTGTCA 1716
1654 TTTCCCGCTGCTTCAGAGCTGAATGACATGTCACCTTGAGTGCATGAGAGTGTCA 1713
1717 CATGGAGTGTATGAGCAGCATATCAAGGCTTCTGTGACATGAGGCGCAAAACCA 1776
1714 CATGGAGTGTATGAGCAGCATATCAAGGCTTCTGTGACATGAGGCGCAAAACCA 1773
1777 GTAAACAGCACCCTTCTTGGAAAGGAGATCAAAAGCTTTGGGCGTGTATGAAAAGACT 1836
1774 GTAAACAGCACCCTTCTTGGAAAGGAGATCAAAAGCTTTGGGCGTGTATGAAAAGACT 1833
1837 TTTTCAATGATACCAATTAACTGCGCTGGAAAGGCGATAGGTGGG 1883
1834 TTTTCAATGATACCAATTAACTGCGCTGGAAAGGCGATAGGTGGG 1880

RESULT 12
AY125488
LOCUS
DEFINITION
Homo sapiens death effector domain-containing DNA-binding 2
AY125488
VERSION
AY125488.1 GI:22475163
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2012)
Lee J.C., Schickling O., Steghn, A.H., Oshima, R.G., Dinsdale, D.,
Chen, G.M. and Peter, M.E.
DDEDD regulates degradation of intermediate filaments during
apoptosis
J. Cell Biol. 158 (6), 1051-1066 (2002)
22202088
JOURNAL
MEDLINE
PUBMED
12235123
REFERENCE
2 (bases 1 to 2012)
Peter, M.E.
Direct Submission
Submitted (21-JUN-2002) Ben May Cancer Institute, University of
Chicago, 924 E. 57th Street, R112, Chicago, IL 60637, USA
Location/Qualifiers

REFERENCE

1..2012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2012
/gene="DDEDD2"
84..1049
/gene="DDEDD2"
/codon_start=1
/product="death effector domain-containing DNA-binding 2
protein"
/protein_id="AAH95240.1"
/db_xref="GI:22475164"

CDS

1..2012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2012
/gene="DDEDD2"
84..1049
/gene="DDEDD2"
/codon_start=1
/product="death effector domain-containing DNA-binding 2
protein"
/protein_id="AAH95240.1"
/db_xref="GI:22475164"

FEATURES
source

Query Match 37.5%; Score 707; DB 9; Length 2012;

ORIGIN

1..2012
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2012
/gene="DDEDD2"
84..1049
/gene="DDEDD2"
/codon_start=1
/product="death effector domain-containing DNA-binding 2
protein"
/protein_id="AAH95240.1"
/db_xref="GI:22475164"

[illegible]

AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	4 (bases 1 to 95663)
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE	Direct Submission
JOURNAL	Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On Jul 13, 2002 this sequence version replaced gi:21637454. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.sbgc.stanford.edu
FEATURES	Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 0.1. NOTE: Small insert shatter library only 85964-86181. 226bp single subclone 9479-9697.
SOURCE	Location/Qualifiers 1..95663 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="19" /clone="CFC-378H22" join(9479..9697,85964..86181) /note="NOTE: Small insert shatter library only 85964-86181. 226bp single subclone 9479-9697."
ORIGIN	
Query Match	37.5%; Score 707; DB 9; Length 95663;
Best Local Similarity	100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 707; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1177 CCCCTTGACAGCCCTCCACAGAGATGTGGGCTCTGAGGCTTAAACATTTCACGCTGAG 1236
Db	90460 CCCCTTGACAGCCCTCCACAGAGATGTGGGCTCTGAGGCTTAAACATTTCACGCTGAG 90400
QY	1237 TTTCCTTCCAGACTCTCTCTTACCCCAAGTGTGCCCTTACGCTCCGAGGCGGGGGC 1296
Db	90400 TTTCCTTCCAGACTCTCTCTTACCCCAAGTGTGCCCTTACGCTCCGAGGCGGGGGC 90340
QY	1297 TGGGCGGTATCTCAGAGGAGGAGGAGCAGCTACACCTCAACCAAGGCCCCCTGAC 1356
Db	90340 TGGGCGGTATCTCAGAGGAGGAGGAGCAGCTACACCTCAACCAAGGCCCCCTGAC 90280
QY	1357 ATTGTATCTGTGATCTTGTGGGCTCTGTGCACTGTCAAGGTGACACACTCGCTCATGTC 1416
Db	90280 ATTGTATCTGTGATCTTGTGGGCTCTGTGCACTGTCAAGGTGACACACTCGCTCATGTC 90220
QY	1417 ACACTGCCCCCTGCTGAGATTTCTCCCTGGGCTCTGTGCCCTTGCTGCTTCCAGACACAC 1476
Db	90220 ACACTGCCCCCTGCTGAGATTTCTCCCTGGGCTCTGTGCCCTTGCTGCTTCCAGACACAC 90160
QY	1477 TTCTTTGGGCGCTAAGGGCTTCTCTCTCAGACCTCTTAATTGACCAACAACACTGGGGCT 1536
Db	90160 TTCTTTGGGCGCTAAGGGCTTCTCTCTCAGACCTCTTAATTGACCAACAACACTGGGGCT 90100
QY	1537 TCAGCCACATCAGTGGGCACTGAGGCTGGGGTGCACATGGGGGCTCTCTCACTTGGCCAC 1596
Db	90100 TCAGCCACATCAGTGGGCACTGAGGCTGGGGTGCACATGGGGGCTCTCTCACTTGGCCAC 90040
QY	1597 ACAATCTCAGCCAGCAGGGCCCTGCCCAGCTCAAATTACAGACTGACTCTCTTACC 1656
Db	90040 ACAATCTCAGCCAGCAGGGCCCTGCCCAGCTCAAATTACAGACTGACTCTCTTACC 89980
QY	1657 TTCCCCCTCTGCTGTCCAGAGCTGAACATGACTTTCGACTTGGATGTTCACCTGGAGGTCA 1716
Db	89980 TTCCCCCTCTGCTGTCCAGAGCTGAACATGACTTTCGACTTGGATGTTCACCTGGAGGTCA 89920
QY	1717 CATGGAGGTGTATGGAGAGCATATACCAAGGCTACTGTGTCACATGGGGCCCAAAACA 1776
Db	89920 CATGGAGGTGTATGGAGAGCATATACCAAGGCTACTGTGTCACATGGGGCCCAAAACA 89860

QY 1777 GTAACGACCACCTTCTTGGAAGGGAATGCAAGGCTTTGGGGGTATGTAAGAAAGACT 1836
DB 89860 GTAACGACCACCTTCTTGGAAGGGAATGCAAGGCTTTGGGGGTATGTAAGAAAGACT 89801
QY 1837 TTTCACAAATGATACCAATTAACCTGCGCTGGAAAGGCGATAGTGGG 1883
DB 89800 TTTCACAAATGATACCAATTAACCTGCGCTGGAAAGGCGATAGTGGG 89754

RESULT 14
AF457575
LOCUS AF457575 981 bp mRNA linear PRI 10-APR-2002
DEFINITION Homo sapiens death effector domain-containing protein FLAME-3
ACCESSION AF457575
KEYWORDS AF457575.1 GI:20126793
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 981) Zhan, Y., Hegde, R., Srinivasula, S.M., Fernandes-Alnemri, T. and
Alnemri, E.S.
TITLE Death effector domain-containing proteins DEDD and FLAME-3 form
nuclear complexes with the TrifC102 subunit of human transcription
factor IIC
JOURNAL Cell Death Differ. 9 (4), 439-447 (2002)
MEDLINE 21961615
PUBMED 11965497
REFERENCE 2 (bases 1 to 981)
AUTHORS Alnemri, E.S.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2001) Microbiology and Immunology, Thomas
Jefferson University, 233 S. 10th Street, Philadelphia, PA 19107,
USA

FEATURES
source location/Qualifiers
1..981
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.2"
1..981
/gene="FLAME-3"
1..981
/gene="FLAME-3"
/codon_start=1
/product="death effector domain-containing protein
FLAME-3"
/protein_id="PAM10835.1"
/db_xref="GI:20126794"
/translation="MALSGSTPAPCWEDECLDYGNLSINRMFEVVGQLTCELEL
LARLLRPPVAGGGLARARGLLELLERGGQDESNRLILGLLRLVLRHLLPHL
ARRRRPVSPEERYSYGTSKRGGSCRRGSSSANGOCOMETGSPTRKORS
RGRPSGABRRRRGAPAAPQOOSGPAPSSSEGVTCIDIRLRAVRAEYRHHGPALEQVVA
SRPPQALRDLDPFGATATLRSDRLGSSVYCDIKFSLSTLDAFMKGYLSEGLIQALR
GVFLTEALRAVGRBEAVRLVSYDEADYEAQRRLIMEEGGRPTES"

ORIGIN
Query Match 35.8%; Score 675; DB 9; Length 981;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 975; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 124 ATGGCGCTATTCGGGCTGACCCCGGCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 183
DB 1 ATGGCGCTATTCGGGCTGACCCCGGCGCTGCTGGAGAGAGATGAGTGGCTGACTAC 60
QY 184 TAGGGAGTGTGTCGCTTCAACCGATGTTGAGAGTGTGGCGGGCAATGACCGACTGC 243
DB 61 TAGGGAGTGTGTCGCTTCAACCGATGTTGAGAGTGTGGCGGGCAATGACCGACTGC 120

QY 244 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCCGCGGAGGCTTAGCC 303
DB 121 GAGCTGAGCTCTGCGCTTTCTGCTGATGAGGCTCTGCGCCGCGGAGGCTTAGCC 180
QY 304 CGGGCCCGCAGCGGCTTAGGCTCTGCTGAGCTGGAAGCGCGGCGGAGCTGCGGAG 363
DB 181 CGGGCCCGCAGCGGCTTAGGCTCTGCTGAGCTGGAAGCGCGGCGGAGCTGCGGAG 240
QY 364 AGCAACCTGCGGCTGCTGGGGCAACTCTGCGGCTGCGGCGGAGGCTGCTGCGG 423
DB 241 AGCAACCTGCGGCTGCTGGGGCAACTCTGCGGCTGCGGCGGAGGCTGCTGCGG 300
QY 424 CACCTGCGCGCAAGCGCGCGCGCGCAAGTCTTCCAGAACGCTATAGCACTCC 483
DB 301 CACCTGCGCGCAAGCGCGCGCGCGCAAGTCTTCCAGAACGCTATAGCACTCC 360
QY 484 AGCTCTTCAAAAGAGACAGAGGCTAGTGCCTGCGCTGCGGAGTCAAGAGTTCTGCA 543
DB 361 AGCTCTTCAAAAGAGACAGAGGCTAGTGCCTGCGCTGCGGAGTCAAGAGTTCTGCA 420
QY 544 AATTCTGAGAGGGTCAAGTGGAGAGACAGGCTCCCGCCCAAGCAAGCGGCGGAGT 603
DB 421 AATTCTGAGAGGGTCAAGTGGAGAGACAGGCTCCCGCCCAAGCAAGCGGCGGAGT 480
QY 604 CGGGCCCGCGCAGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCAGCGCACCCAG 663
DB 481 CGGGCCCGCGCAGTGTGTGTCAGACGCGCGCGAGAGAGGGGCCAGCGCACCCAG 540
QY 664 CAGCAGTCAAGGCGCGCAGACCTTCTCTGAAAGCAAAGTACCTTGAACATCCGCTC 723
DB 541 CAGCAGTCAAGGCGCGCAGACCTTCTCTGAAAGCAAAGTACCTTGAACATCCGCTC 600
QY 724 CGGGTTGAGACAGATCTGCGGAGCAATGGGCGAGGCTTGGAGAGAGGCGGTGCGA 783
DB 601 CGGGTTGAGACAGATCTGCGGAGCAATGGGCGAGGCTTGGAGAGAGGCGGTGCGA 660
QY 784 CGGCCCCAGCGCTGCGCGCGCAGCTGAGCGTGTGGGAGGCGCACGAGTGTGCGC 843
DB 661 CGGCCCCAGCGCTGCGCGCGCAGCTGAGCGTGTGGGAGGCGCACGAGTGTGCGC 720
QY 844 TCAGAGGACCTGCGCTGTGTGTTGTGACATCAAGTTCTCAAGCTCTCTATCTGAC 903
DB 721 TCAGAGGACCTGCGCTGTGTGTTGTGACATCAAGTTCTCAAGCTCTCTATCTGAC 780
QY 904 GCCCTCGGGGCGACATCACTGAGTGGCGGCTGCGAGGCGCGGCGGCGTTCCTG 963
DB 781 GCCCTCGGGGCGACATCACTGAGTGGCGGCTGCGAGGCGCGGCGGCGTTCCTG 840
QY 964 ACTAGGCGCTGCGAGAGGCTGTGTGGCGCGGAGGCTGTGCGCTGCTGATGTGAT 1023
DB 841 ACTAGGCGCTGCGAGAGGCTGTGTGGCGCGGAGGCTGTGCGCTGCTGATGTGAT 900
QY 1024 GAGGCTGACTATAGAGCTGTGCGCGCGCGCTGTGCTGATGAGAGAGAGGGGCGG 1083
DB 901 GAGGCTGACTATAGAGCTGTGCGCGCGCGCTGTGCTGATGAGAGAGAGGGGCGG 960
QY 1084 CGCCCGACAGAGGCGCTCTGGA 1104
DB 961 CGCCCGACAGAGGCGCTCTGGA 981

RESULT 15
BD149917
LOCUS BD149917 626 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD149917
VERSION BD149917.1 GI:27855675
KEYWORDS JP 2002191363-A/4760.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 626)

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 4760 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/4760
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAMORU
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10', C12P21/02, C12Q1/68//C12P21/08 G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..626
/organism='Homo sapiens (human)'.
location/Qualifiers
1..626
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 22.4%; Score 421; DB 6; Length 626;
Best Local Similarity 99.8%; Pred. No.3e-210;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 37 AACAACTGCTGCGCGCGCTGCGCCAGAGATCATCGACGCCAGAACTTGCGCGGTTCTG 96
Db 1 AACAACTGCTGCGCGCGCTGCGCCAGAGATCATCGACGCCAGAACTTGCGCGGTTCTG 60

QY 97 AGCTTTGTTCCGCTCTCCCTCCCGGGGAATGGCGCTATCCGGGTTCGACCCCGCGCCGTGC 156
Db 61 AGCTTTGTTCCGCTCTCCCTCCCGGGGAATGGCGCTATCCGGGTTCGACCCCGCGCCGTGC 120

QY 157 TGGAGAGAGATGAGAGGCTGAGCTAGCTAGGAGATGCTGTGCTTCAACCGTATGTTCCGAG 216
Db 121 TGGAGAGAGATGAGAGGCTGAGCTAGCTAGGAGATGCTGTGCTTCAACCGTATGTTCCGAG 180

QY 217 GTGCTGCGCGCGCACTGACCGAGTGCAGCTGAGCTTCTGCGCTTTCTGCTGATGAG 276
Db 181 GTGCTGCGCGCGCACTGACCGAGTGCAGCTGAGCTTCTGCGCTTTCTGCTGATGAG 240

QY 277 GCTCTGCGCGCGCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGCTCTGCTGAG 336
Db 241 GCTCTGCGCGCGCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGCTCTGCTGAG 300

QY 337 CTGAGAGCGCGCGCGGAGTGCAGGAGAGCAACTGCGCTGCTGCGGGCAACTCTGCGCG 396
Db 301 CTGAGAGCGCGCGCGGAGTGCAGGAGAGCAACTGCGCTGCTGCGGGCAACTCTGCGCG 360

QY 397 GTGCTGCGCGCGCGCACTGCTGCTGCGCACTTGGCGCGCAAGCGCGCGCGCAAGTGTCT 456
Db 361 GTGCTGCGCGCGCGCACTGCTGCTGCGCACTTGGCGCGCAAGCGCGCGCGCAAGTGTCT 420

QY 457 CCAGAACTGATAGCTATGAGCACTTCAAGCTTTCAAAAGAGACAGAGGTA 508
Db 421 CCAGAACTGATAGCTATGAGCACTTCAAGCTTTCAAAAGAGACAGAGGTA 472

RESULT 16
LOCUS AX869855 626 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 4760 from Patent Epi074617.
ACCESSION AX869855
VERSION AX869855.1 GI:40024718
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 4760 07-FEB-2001;
Research Association for Biotechnology (JP)
Location/Qualifiers
FEATURES
source 1..626
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 22.4%; Score 421; DB 6; Length 626;
Best Local Similarity 99.8%; Pred. No.3e-210;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 37 AACAACTGCTGCGCGCGCTGCGCCAGAGATCATCGACGCCAGAACTTGCGCGGTTCTG 96
Db 1 AACAACTGCTGCGCGCGCTGCGCCAGAGATCATCGACGCCAGAACTTGCGCGGTTCTG 60

QY 97 AGCTTTGTTCCGCTCTCCCTCCCGGGGAATGGCGCTATCCGGGTTCGACCCCGCGCCGTGC 156
Db 61 AGCTTTGTTCCGCTCTCCCTCCCGGGGAATGGCGCTATCCGGGTTCGACCCCGCGCCGTGC 120

QY 157 TGGAGAGAGATGAGAGGCTGAGCTAGCTAGGAGATGCTGTGCTTCAACCGTATGTTCCGAG 216
Db 121 TGGAGAGAGATGAGAGGCTGAGCTAGCTAGGAGATGCTGTGCTTCAACCGTATGTTCCGAG 180

QY 217 GTGCTGCGCGCGCACTGACCGAGTGCAGCTGAGCTTCTGCGCTTTCTGCTGATGAG 276
Db 181 GTGCTGCGCGCGCACTGACCGAGTGCAGCTGAGCTTCTGCGCTTTCTGCTGATGAG 240

QY 277 GCTCTGCGCGCGCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGCTCTGCTGAG 336
Db 241 GCTCTGCGCGCGCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGCTCTGCTGAG 300

QY 337 CTGAGAGCGCGCGCGGAGTGCAGGAGAGCAACTGCGCTGCTGCGGGCAACTCTGCGCG 396
Db 301 CTGAGAGCGCGCGCGGAGTGCAGGAGAGCAACTGCGCTGCTGCGGGCAACTCTGCGCG 360

QY 397 GTGCTGCGCGCGCGCACTGCTGCTGCGCACTTGGCGCGCAAGCGCGCGCGCAAGTGTCT 456
Db 361 GTGCTGCGCGCGCGCACTGCTGCTGCGCACTTGGCGCGCAAGCGCGCGCGCAAGTGTCT 420

QY 457 CCAGAACTGATAGCTATGAGCACTTCAAGCTTTCAAAAGAGACAGAGGTA 508
Db 421 CCAGAACTGATAGCTATGAGCACTTCAAGCTTTCAAAAGAGACAGAGGTA 472

RESULT 17
LOCUS AR339362 1067 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 853 from patent US 6569662.
ACCESSION AR339362
VERSION AR339362.1 GI:33726219
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1067)
AUTHORS Tang,Y.T., Zhou,P. and Drmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6569662-A 853 27-MAY-2003;
Location/Qualifiers
FEATURES
source 1..1067
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 19.5%; Score 368; DB 6; Length 1067;
Best Local Similarity 99.4%; Pred. No. 2.8e-182;
Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 37 AACAGGCGTGGCGCGCGTGGCCAGAGATCATGGACCGCAGAAATCTGGCCGGTCTG 96
DB 64 AACAGGCGTGGCGCGCGTGGCCAGAGATCATGGACCGCAGAAATCTGGCCGGTCTG 123
QY 97 AGCTTTGTTCCGCTCCCTCCCGGGAATGACGCTATCCGGGTGACCCCGGCGCGTGC 156
DB 124 AGCTTTGTTCCGCTCCCTCCCGGGAATGACGCTATCCGGGTGACCCCGGCGCGTGC 183
QY 157 TGGAGAGAGATGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 216
DB 184 TGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243
QY 217 GTGGTGGGCGGCGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 276
DB 244 GTGGTGGGCGGCGCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
QY 277 GCTCTGGGCGGCGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGAC 336
DB 304 GCTCTGGGCGGCGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGAC 363
QY 337 CTGGAGCGCGCGGCGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGAC 396
DB 364 CTGGAGCGCGCGGCGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGAC 423
QY 397 GTGGTGGGCGGCGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGAC 456
DB 424 GTGGTGGGCGGCGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGAC 483
QY 457 CCAGAACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 516
DB 484 CCAGAACGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
QY 517 CGCGGTGGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGACCGG 557
DB 544 CGCGGTGGGCGGCTTATGACCGGCGGCGGCTTATGACCGGCGGCTTATGACCGG 584
```

RESULT 18
AC006486/c 111084 bp DNA linear PRI 02-FEB-1999
LOCUS Homo sapiens chromosome 19, BAC CIT-B-147823 (BC85722), complete
DEFINITION
AC006486
VERSION AC006486
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Lamerdin,J.E., McCready,P.M., Skowronski,E., Viswanathan,V.,
Burchhart-Schulz,K., Gordon,L., Dias,J., Ramirez,M., Stiliwgen,S.,
Pan,H., Velasco,N., Do,L., Regala,W., Terry,A., Garnes,J.,
Liu,S., Actix,C., Andreise,T., Frankheim,M., Amico-Keller,G.,
Coetfield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G.,
Kornmiller,B., Arellano,A., Sanders,C., Ow,D., Nolan,M., Trong,S.,
Kobayashi,A., Olsen,A.S. and Carrano,A.V.
TITLE Sequence analysis of a 3.2 Mb region in 19q13.2 between CYP2F1 and
D19S178
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 111084)
AUTHORS Lamerdin,J.E.
JOURNAL Direct Submission
COMMENT Submitted (02-FEB-1999) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
MAP and sequence oriented from q centromere to telomere, BAC 85722
(CIT-B-147823) is separated from BAC 48537 (CIT-B-50E14) to the
right by a gap of approximately 20 Kb. There is currently no

neighboring clone centromeric of BAC 85722 in the sequencing queue.
Additional chromosome 19 map and sequence information are available
from: <http://www-bio.1lnl.gov/genome/genome.html>.
Location/Qualifiers

```
FEATURES
    source
        1..111084
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="19"
            /map="19q13.2 between CYP2F1 and D19S178"
            /clone="CIT-B-147823"
            /note="LNL clone name: BC85722"
    repeat_region
        /rpt_family="LINE2"
        complement(707..805)
        /rpt_family="MIR"
        811..1115
        /rpt_family="AluSx"
        complement(1121..1244)
        /rpt_family="MIR"
        complement(1561..1695)
        /note="predicted exon, program: gra12exons human_1.3,
        frame: 0, quality: good, score: 52.000-(1656..1558) DDS
        similarity to AA725300 a112b07.81 Soares parathyroid tumor
        NblPA Homo sapiens cDNA clone 1342549 3'; (466..369); 98%
        identity."
        complement(2247..2549)
        /rpt_family="AluSg"
        2639..2944
        /rpt_family="AluSx"
        3015..3255
        /rpt_family="AluSx"
        3257..3340
        /rpt_family="AluSx"
        3611..3912
        /rpt_family="AluSx"
        3913..3934
        /rpt_family="AluSx"
        3939..4228
        /rpt_family="AluSx"
        complement(4241..4311)
        /rpt_family="AluSx"
        complement(4322..4618)
        /rpt_family="AluSx"
        complement(4623..4743)
        /rpt_family="FLM_C"
        4832..5013
        /rpt_family="L1MC4"
        5131..6020
        /rpt_family="L1MC4"
        6028..6328
        /rpt_family="AluSx"
        6341..6639
        /rpt_family="AluSx"
        complement(6648..6915)
        /note="BLASTN similarity to X97703 (1..267); match: 0.98,
        score: 1.1e-95; database searched: nt; H.sapiens DNA for
        highly polymorphic microsatellite marker"
        6856..6886
        /rpt_family="CA_n"
        complement(6988..7107)
        /note="predicted exon, program: gra12exons human_1.3,
        frame: 0, quality: excellent, score: 88.000"
        7851..7908
        /rpt_family="LINE2"
        8169..8329
        /rpt_family="MIR"
        complement(8535..8862)
        /note="predicted exon, program: gra12exons human_1.3,
        frame: 0, quality: excellent, score: 77.000"
        8543..8765
        /note="DGS similarity to AA511237 vhs604.r1 Soares mouse
        mammary gland NDMMG Mus musculus cDNA clone 890983 5';
```

	repeat_region	22340..22405	/rpt_family="(GGA)n"	complement(join(22649..22722,23912..24004,24351..24537,24965..25060,25141..25238,26208..26314,26403..26533,28461..28571,28713..28796,31810..31997,34038..34320))
CDS			/gene="GSK3A"	
			/note="GLYCOCEN SYNTHASE KINASE-3 ALPHA; GSK-3 ALPHA"	
			/codon_start=1	
			/product="K3A_HUMAN"	
			/protein_id="ABD11986.1"	
			/db_xref="GI:4210499"	
			/translation="MSGGGPSSGGPGSGSPARTSFAEPGGGGGGGGPPGSASGPGGTGGGAGSVAMGGGVASSGGGPGSGSGGPPAGTSFPFGVKLRBDSKVTTVATLGGGPERSOEVAATTDIKVINGSGFVVYAPARLETRELAIKVLDDXPKNELOIMRKLDHCNIVRLRIFFYSSEKKDELYLVLEVPEPTVRYVARHPFKALCTPI LVYKNVYQLFRSLAYTHSQGVCHRIDPLPNMLVDPTVALKLCDPSAQLYRGEN VSAICSRYPARELPFIQATITYSIDWSAGCVLAELLGQPIFPGSGVDLVEIIL VLEPTREQIRKMNPMPYEKEFPQIKHWPMTKFSTPREBALCSLSLEYPSSEL SPLEACHSFDELRCIGTOLPNNRPPLPLEFNFSAGBLSIQSPINAILIPPHRSFAG TTTLTPESSQALTETPTPSSDMWGSDATFTLTNSS"	
	repeat_region	22935..23225	/rpt_family="Alusg"	
	repeat_region	23355..23646	/rpt_family="Alusx"	
			/rpt_family="Alusx"	
Query Match	Best Local Similarity	16.8%; Score 316; DB 9; Length 111084;		
Matches	366; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Oy	85	GGCCGGGTTGAGACTTGTCGCCGCTCCCGCCCCCGGAAATGGCGCATCCGGGTGACC	144	
Db	8901	GGCCGGGTTCTGACCTTGTCGCCCTCCTCCCGGAAATGGCGCATCCGGGTGACC	8842	
Oy	145	CCGGCCCCGTCCTGGAGAGAAGATAGTCCTGACTACTACGGAGTCTGTGCCTTCA	204	
Db	8841	CCGGCCCCGTCCTGGAGAGAAGATAGTCCTGACTACTACGGAGTCTGTGCCTTCA	8782	
Oy	205	CGTATGTTCCAGGTGTGTGGCGGGCACTGACCGAGTGCAGCTTGAGGCTTCCTGGCCTTT	264	
Db	8781	CGTATGTTCCAGGTGTGTGGCGGGCACTGACCGAGTGCAGCTTGAGGCTTCCTGGCCTTT	8722	
Oy	265	CTGCTGATGATGAGCTCTCGGGGCGCGGAGAGGCTTACCCGGGGCCGACAGCGGCTTAG	324	
Db	8721	CTGCTGATGATGAGCTCTCGGGGCGCGGAGAGGCTTACCCGGGGCCGACAGCGGCTTAG	8662	
Oy	325	CTCCTGTGAGAGCTGAGCGCGCGGCGCAGTGCGGCGAGACACTGCGGCTTCTGGGG	384	
Db	8661	CTCCTGTGAGAGCTGAGCGCGCGGCGCAGTGCGGCGAGACACTGCGGCTTCTGGGG	8602	
Oy	385	CAACTCTCTGCGCGCTGTGGCCCGCACGACCTGCTGCCGACCTGGCGCGACAGCGCGC	444	
Db	8601	CAACTCTCTGCGCGCTGTGGCCCGCACGACCTGCTGCCGACCTGGCGCGACAGCGCGC	8542	
Oy	445	CGGCGAG 451		
Db	8541	CGGCGAG 8535		
RESULT 19				
LOCUS	AX136449	603 bp	DNA	linear PAT 30-MAY-2001
DEFINITION	Sequence 371 from Patent EP1067182.			
ACCESSION	AX136449			
KEYWORDS	AX136449.1 GI:14272853			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 Oca,T., Isegai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.			
TITLE	Secretory protein or membrane protein			

JOURNAL Patient: EP 1067182-A 371 10-JAN-2001;
 FEATURES Helix Research Institute (JP)
 source Location/Qualifiers
 1..603
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

Query Match 14.2%; Score 268; DB 6; Length 603;
 Best Local Similarity 100.0%; Pred. No. 2.1e-129;
 Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CCAAAACAAATGCTGCGCCGCTGCGCCAGAGTCAATCGACGACGAGATCTGCGCGGT 92
 DB 32 CCAAAACAAATGCTGCGCCGCTGCGCCAGAGTCAATCGACGACGAGATCTGCGCGGT 91
 QY 93 TCTGAGCTTGTTCGCGCTCCCTCCCGGGAATGCGGCTATCCGGTTCGACCCCGGCCCC 152
 DB 92 TCTGAGCTTGTTCGCGCTCCCTCCCGGGAATGCGGCTATCCGGTTCGACCCCGGCCCC 151
 QY 153 GTGCTGGAGAGAGATAGTGCCTGGACTACTAGGAGATGCTGCTTCAACCGTATGTT 212
 DB 152 GTGCTGGAGAGAGATAGTGCCTGGACTACTAGGAGATGCTGCTTCAACCGTATGTT 211
 QY 213 CGAGGTGCTGGCGCGGCACTGACCGAGTGCAGAGCTGAGCTCTGGCCTTTCTGTGGA 272
 DB 212 CGAGGTGCTGGCGCGGCACTGACCGAGTGCAGAGCTGAGCTCTGGCCTTTCTGTGGA 271
 QY 273 TGAAGCTCTGCGCGCGCGGAGGCTTA 300
 DB 272 TGAAGCTCTGCGCGCGCGGAGGCTTA 299

RESULT 20
 LOCUS BD123689 603 bp DNA linear PAT 18-SEP-2002
 DEFINITION Secretory protein or membrane protein.
 ACCESSION BD123689
 VERSION BD123689.1 GI:23218634
 KEYWORDS JP 2002017376-A/198.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 603)
 AUTHORS Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and
 Hayashi,K.
 TITLE Secretory protein or membrane protein
 JOURNAL Patient: JP 2002017376-A 198 22-JAN-2002;
 HELIX RESEARCH INSTITUTE
 COMMENT OS Homo sapiens (human)
 PN JP 2002017376-A/198
 PD 22-JAN-2002
 PF 07-JUL-2000 JP 2000253173
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,YURI KAWAI,TOMOYASU
 PI SUGIYAMA,
 PI KOJI HAYASHI
 PC

PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
 10,
 PC C12P21/02,C12Q1/68//C12P21/08,C12N15/00,C12N5/00 CC
 Secretory protein or membrane protein
 FH Key Location/Qualifiers
 FT source 1..603
 /organism="Homo sapiens (human)"

FEATURES
 source

1..603
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 14.2%; Score 268; DB 6; Length 603;
 Best Local Similarity 100.0%; Pred. No. 2.1e-129;
 Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CCAAAACAAATGCTGCGCCGCTGCGCCAGAGTCAATCGACGACGAGATCTGCGCGGT 92
 DB 32 CCAAAACAAATGCTGCGCCGCTGCGCCAGAGTCAATCGACGACGAGATCTGCGCGGT 91
 QY 93 TCTGAGCTTGTTCGCGCTCCCTCCCGGGAATGCGGCTATCCGGTTCGACCCCGGCCCC 152
 DB 92 TCTGAGCTTGTTCGCGCTCCCTCCCGGGAATGCGGCTATCCGGTTCGACCCCGGCCCC 151
 QY 153 GTGCTGGAGAGAGATAGTGCCTGGACTACTAGGAGATGCTGCTTCAACCGTATGTT 212
 DB 152 GTGCTGGAGAGAGATAGTGCCTGGACTACTAGGAGATGCTGCTTCAACCGTATGTT 211
 QY 213 CGAGGTGCTGGCGCGGCACTGACCGAGTGCAGAGCTGAGCTCTGGCCTTTCTGTGGA 272
 DB 212 CGAGGTGCTGGCGCGGCACTGACCGAGTGCAGAGCTGAGCTCTGGCCTTTCTGTGGA 271
 QY 273 TGAAGCTCTGCGCGCGCGGAGGCTTA 300
 DB 272 TGAAGCTCTGCGCGCGCGGAGGCTTA 299

RESULT 21
 LOCUS AX431298 303 bp DNA linear PAT 28-JUN-2002
 DEFINITION Sequence 7 from Patent WO240680.
 ACCESSION AX431298
 VERSION AX431298.1 GI:21656167
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 AUTHORS Pawlowski,K., Fiorentino,L., Godzik,A., Lee,S.H., Reed,J.C.,
 Roth,W. and Stenner-Liessen,F.
 TITLE Novel death domain proteins
 JOURNAL Patient: WO 0240680-A 7 23-MAY-2002;
 BURNHAM INST (US)
 FEATURES Location/Qualifiers
 source 1..303
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 <1..>303
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37275.1"
 /db_xref="GI:21656168"
 /translation="WEDECLDYGMKSLHMFVEVVGQITGCELEILAFLEBAPGA
 AGGLARARSGLELLELRGQCDENLRLLGQLRLVLAHDLPLHARRRRPVP"

CDS
 1..303
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 <1..>303
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD37275.1"
 /db_xref="GI:21656168"
 /translation="WEDECLDYGMKSLHMFVEVVGQITGCELEILAFLEBAPGA
 AGGLARARSGLELLELRGQCDENLRLLGQLRLVLAHDLPLHARRRRPVP"

ORIGIN

Query Match 13.4%; Score 252; DB 6; Length 303;
 Best Local Similarity 99.7%; Pred. No. 6.4e-121;
 Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 157 TGGAGAGAGATAGTGCCTGGACTACTAGGAGATGCTGCTTCAACCGTATGTTGAG 216
 DB 1 TGGAGAGAGATAGTGCCTGGACTACTAGGAGATGCTGCTTCAACCGTATGTTGAG 60
 QY 217 GTGCTGGCGCGGCACTGACCGAGTGCAGAGCTGAGCTCTGGCCTTTCTGTGATGAG 276
 DB 61 GTGCTGGCGCGGCACTGACCGAGTGCAGAGCTGAGCTCTGGCCTTTCTGTGATGAG 120
 QY 277 GCTCCTGGCGCGCGCGGAGGCTTACGCGGCGCGGAGCGGCGCTTAGAGCTCTGCTGAG 336
 DB 121 GCTCCTGGCGCGCGCGGAGGCTTACGCGGCGCGGAGCGGCGCTTAGAGCTCTGCTGAG 180
 QY 337 CTGAGAGCGCGCGGAGTGCAGAGCAACTGCGGCTGCTGGGCAACTCTGCGC 396

Db	181	CTGAGAGCGCCGCGGACAGTGGCGAGAGACCACTGGCGCTGCTGGGCGCACTCTTGCGC	240
Qy	397	GTGCTGGCCCGCCACGACCTGCTGCGCAGCTTGCGCGCAAGCGCGCCGCGCAGTGTCT	456
Db	241	GTGCTGGCCCGCCACGACCTGCTGCGCAGCTTGCGCGCAAGCGCGCCGCGCAGTGTCT	300
Qy	457	CCA 459	
Db	301	CCA 303	
RESULT 22			
LOCUS	CQ460291	441 bp	DNA
DEFINITION	Sequence 3069 from Patent WO0192581.	linear	PAT 30-JAN-2004
ACCESSION	CQ460291		
VERSION	CQ460291.1	GI:41425910	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE			
AUTHORS	Algate, P.A., Harlocker, S.L. and Jones, R.		
TITLE	Compositions and methods for the therapy and diagnosis of		
	ovariocancer		
JOURNAL	Patent: WO 0192581-A 3069 06-DEC-2001;		
	CORIXA CORPORATION (US)		
FEATURES			
source	location/Qualifiers		
	1..441		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	12.6%; Score 237; DB 6; Length 441;		
Best Local Similarity	100.0%; Pred. No. 5.2e-113;		
Matches 237; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
Qy	1209	TCTGAGGCGCTAAACCATTTCCAGCTGAGTTCCTTCCAGACTCCTCTTACCCTCCAGGTG	1268
Db	199	TCTGAGGCGCTAAACCATTTCCAGCTGAGTTCCTTCCAGACTCCTCTTACCCTCCAGGTG	258
Qy	1269	TGCCCCCTTACCCCTCCCGAGGCGGGGCGCTGGGCCCTGATCTCAGAAAGGAGGGGCACAGC	1328
Db	259	TGCCCCCTTACCCCTCCCGAGGCGGGGCGCTGGGCCCTGATCTCAGAAAGGAGGGGCACAGC	318
Qy	1329	TACACACTCACAAAGGCCCCCTTGACATTTGATCTGTGATCTTTGGGCTGTGTGCACTG	1388
Db	319	TACACACTCACAAAGGCCCCCTTGACATTTGATCTGTGATCTTTGGGCTGTGTGCACTG	378
Qy	1389	TCACAGGTGCACACACTGCTCATGCTCACACTGCCCTGTGCTGAGATTTTCCCTGAGG	1445
Db	379	TCACAGGTGCACACACTGCTCATGCTCACACTGCCCTGTGCTGAGATTTTCCCTGAGG	435
RESULT 23			
LOCUS	G20341	222 bp	DNA
DEFINITION	human STS A005M44, sequence tagged site.	linear	STS 24-JUL-1996
ACCESSION	G20341		
VERSION	G20341.1	GI:1340678	
KEYWORDS	STS; STS sequence; primer; sequence tagged site.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE			
AUTHORS	1 (bases 1 to 222)		
JOURNAL	Adams, M.D.		
COMMENT	Unpublished (1996)		
	Contact: Mark Adams		

```

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Email: mtdadams@igf.org

Primer A: CATTGTGAAAAGTCTTTTC
Primer B: AGCTTCATTTCACAGACCT
STS size: 222
PCR Profile:
    Denaturation: 96C 5min
    Anneal: 54C 30sec
    Extend: 72C 30sec
    Denature: 95C 30sec
    FinalExtend: 72C 5min
    Cycles: 30

Protocol:
    GenomicDNA: 25 ng
    Primer: 0.43 uM each
    dNTPs: 230 uM each
    AmpliTaq: 0.5 units
    TagStarT Ab: 0.5 units
    Total Volume: 10 ul

Buffer:
    Tris-HCl pH8.8: 100 mM
    KCl: 500 mM
    Triton X-100: 1%
    Concentration: 10X

Prepared with primer pairs derived from THC87979: GenBank Accession
Numbers- T50230, T50243, L54156.

FEATURES
    source
        1..222
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            1..222
            1..20
            complement(204..222)

STS
    primer_bind
        primer_bind
            ORIGIN

Query Match
    10.3%; Score 194; DB 11; Length 222;
    Best Local Similarity 100.0%; Pred. No. 2.9e-90;
    Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1625 AGCTTCATTTCACAGACCTGCTCTCTCACTTCCCTGCTGTCACAGAGTGAACAT 1684
Db 222 AGCTTCATTTCACAGACCTGCTCTCTCTCACTTCCCTGCTGTCACAGAGTGAACAT 163
QY 1685 AGACTTGACCTTGGATGTCACTGAGATGTCACTGGGAGTGTATATGCGACATCATACC 1744
Db 162 AGACTTCACCTTGGATGTCACTGAGATGTCACTGGGAGTGTATATGCGACATCATACC 103
QY 1745 AAGGCTACTGTTTCACATGGGAGCCCAAAACACAGTAAACAGGCACCTTCTTGAAAAGGAAA 1804
Db 102 AAGGCTACTGTTTCACATGGGAGCCCAAAACACAGTAAACAGGCACCTTCTTGAAAAGGAAA 43
QY 1805 TGCAAAGGCTTTGG 1818
Db 42 TGCAAAGGCTTTGG 29

RESULT 24
AX364851 1084 bp DNA linear PAT 15-FEB-2002
LOCUS AX364851
DEFINITION Sequence 2 from Patent W00206315.
ACCESSION AX364851
VERSION AX364851.1 GI:18696740
KEYWORDS
SOURCE
    Homo sapiens (human)
    Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```

REFERENCE 1
AUTHORS Mintz,L., Freilich,S. and Bernstein,J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 2 24-JAN-2002;
CompuGen Ltd. (IL)
FEATURES
Source Location/Qualifiers
1. 1084
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 9.9%; Score 187; DB 6; Length 1084;
Best Local Similarity 100.0%; Pred. No. 1.2e-86;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 810 GGAAGTGTGGGAGGAGCCAGAGTGTGCTCAAGGAGCTGGGCTGTGTGTTG 869
DB 764 GGAAGTGTGGGAGGAGCCAGAGTGTGCTCAAGGAGCTGGGCTGTGTGTTG 823
QY 870 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGCACTAAGTGG 929
DB 824 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGCACTAAGTGG 883
QY 930 CGCCCTGCTCAGAGCCCTGCGGGGCTGTCTCTGACTGAGGCTCTGCGAGAGGCTGTGG 989
DB 884 CGCCCTGCTCAGAGCCCTGCGGGGCTGTCTCTGACTGAGGCTCTGCGAGAGGCTGTGG 943
QY 990 CCGGAG 996
DB 944 CCGGAG 950
RESULT 25
AX364852 1106 bp DNA linear PAT 15-FEB-2002
LOCUS Sequence 3 from Patent WO0206315.
DEFINITION AX364852
ACCESSION AX364852
VERSION AX364852.1 GI:18696741
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Mintz,L., Freilich,S. and Bernstein,J.
TITLE Novel nucleic acid and amino acid sequences
JOURNAL Patent: WO 0206315-A 3 24-JAN-2002;
CompuGen Ltd. (IL)
FEATURES
Source Location/Qualifiers
1. 1106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 9.9%; Score 187; DB 6; Length 1106;
Best Local Similarity 100.0%; Pred. No. 1.2e-86;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 810 GGAAGTGTGGGAGGAGCCAGAGTGTGCTCAAGGAGCTGGGCTGTGTGTTG 869
DB 786 GGAAGTGTGGGAGGAGCCAGAGTGTGCTCAAGGAGCTGGGCTGTGTGTTG 845
QY 870 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGCACTAAGTGG 929
DB 846 TGACATCAAGTTCTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGCACTAAGTGG 905
QY 930 CGCCCTGCTCAGAGCCCTGCGGGGCTGTCTCTGACTGAGGCTCTGCGAGAGGCTGTGG 989
DB 906 CGCCCTGCTCAGAGCCCTGCGGGGCTGTCTCTGACTGAGGCTCTGCGAGAGGCTGTGG 965
QY 990 CCGGAG 996

DB 966 CCGGAG 972
RESULT 26
BD154885/c 381 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD154885
ACCESSION BD154885.1 GI:27860643
VERSION BD154885.1 GI:27860643
KEYWORDS JP 2002191363-A/9728.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 381)
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,T.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 9728 09-JUL-2002;
HELIOS RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/9728
PD 09-JUL-2002
PF 28-JUL-2000 JP 2002280990
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
PI SATO,
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
PI KEIICHI NAGAI,TETSUJI OTSUKI
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
C12P21/02,C12Q1/68//C12P21/08,G06P17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH key
FEATURES
Source Location/Qualifiers
FT source 1. 381
location/Qualifiers
1. 381
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 8.8%; Score 165; DB 6; Length 381;
Best Local Similarity 99.5%; Pred. No. 5.9e-75;
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1668 TGTCAGAGCTGAACATAGACTTGCACCTTGATGTCACCTGAGTGCATGAGGAGTGT 1727
DB 216 TGTCAGAGCTGAACATAGACTTGCACCTTGATGTCACCTGAGTGCATGAGGAGTGT 157
QY 1728 TATGCGAGCATATACCAAGGCTTACTGTTGCAATGCGGCCCAAAACAGTAACACGCCA 1787
DB 156 TATGCGAGCATATACCAAGGCTTACTGTTGCAATGCGGCCCAAAACAGTAACACGCCA 97
QY 1788 CTTTCTGGAAGAGGAATGCAAGGCTTTGGGGGTGATGGAAGACCTTTTCAATGA 1847
DB 96 CTTTCTGGAAGAGGAATGCAAGGCTTTGGGGGTGATGGAAGACCTTTTCAATGA 37
QY 1848 TACCAATTAATAGCTCCCTGGAAGAGGAGATAGTGG 1883
DB 36 TACCAATTAATAGCTCCCTGGAAGAGGAGATAGTGG 1
RESULT 27
AX874823/c 381 bp DNA linear PAT 17-DEC-2003
LOCUS AX874823
DEFINITION Sequence 9728 from Patent EP1074617.
ACCESSION AX874823
VERSION AX874823.1 GI:40029582
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isozaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Makamatsu, A., Nagai, K., and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 9728 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source 1.381
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 8.8%; Score 165; DB 6; Length 381;
Best Local Similarity 99.5%; Pred. No. 5.9e-75;
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1668 TGTCCAGAGCTGAACATAGACTTGCACCTTGATGTCACCTGAGTGCATGAGGAGTGT 1727
Db 216 TGTCCAGAGCTGAACATAGACTTGCACCTTGATGTCACCTGAGTGCATGAGGAGTGT 157
QY 1728 TATGGACAGCATACACCAAGGCTTGTGTCACATGGGSCCAAAACAGTAAACAGCCA 1787
Db 156 TATGGACAGCATACACCAAGGCTTGTGTCACATGGGSCCAAAACAGTAAACAGCCA 97
QY 1788 CCTTCTTGAAGAGGATGCAAGGCTTGGGGGTGTGTGAAAGACCTTTTCAATGA 1847
Db 96 CCTTCTTGAAGAGGATGCAAGGCTTGGGGGTGTGTGAAAGACCTTTTCAATGA 37
QY 1848 TACCAATTAACTGCCCTGGAAAGGGCATAGTGGG 1883
Db 36 TACCAATTAACTGCCCTGGAAAGGGCATAGTGGG 1
RESULT 28
AC068283/c 167108 bp DNA linear HTG 30-AUG-2000
LOCUS Homo sapiens chromosome 19 clone RP11-317E13, WORKING DRAFT
DEFINITION
AC068283 167108 bp DNA linear HTG 30-AUG-2000
SEQUENCE 36 unordered pieces.
AC068283
AC068283.4 GI:9945176
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 167108)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 167108)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (30-APR-2000) Genome Sequencing Center, Washington
REFERENCE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
TITLE MO 63108, USA
JOURNAL On Aug 30, 2000 this sequence version replaced gi:9558638.
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0317E13
----- Summary Statistics -----
Sequencing vector: MJ3, 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319

Consensus quality: 146259 bases at least Q40
Consensus quality: 153006 bases at least Q30
Consensus quality: 156230 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 163958; sum-of-contigs
Quality coverage: 3.94 in Q20 bases; agarose-fp
Quality coverage: 4.25 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 36 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1540: contig of 1540 bp in length
1541 1640: gap of unknown length
1641 2783: contig of 1143 bp in length
2784 2883: gap of unknown length
2884 4448: contig of 1565 bp in length
4449 4548: gap of unknown length
4549 6252: contig of 1703 bp in length
6252 6352: gap of unknown length
6352 9081: contig of 2730 bp in length
9082 9181: gap of unknown length
9182 10460: contig of 1279 bp in length
10461 10560: gap of unknown length
10561 12863: contig of 2303 bp in length
12864 12963: gap of unknown length
12964 15336: contig of 2573 bp in length
15337 15537: gap of unknown length
15537 18260: contig of 2624 bp in length
18261 18360: gap of unknown length
18361 20681: contig of 2321 bp in length
20682 20781: gap of unknown length
20782 24034: contig of 3253 bp in length
24035 24135: gap of unknown length
24135 27582: contig of 3447 bp in length
27582 27681: gap of unknown length
27682 30155: contig of 2473 bp in length
30155 30254: gap of unknown length
30255 33662: contig of 3408 bp in length
33663 33762: gap of unknown length
33763 35826: contig of 2064 bp in length
35827 35926: gap of unknown length
35927 39520: contig of 3594 bp in length
39521 39620: gap of unknown length
39621 42496: contig of 2876 bp in length
42497 42596: gap of unknown length
42597 45290: contig of 2694 bp in length
45291 45390: gap of unknown length
45391 50042: contig of 4652 bp in length
50043 50142: gap of unknown length
50143 54382: contig of 4240 bp in length
54383 54482: gap of unknown length
54483 58846: contig of 4364 bp in length
58847 58946: gap of unknown length
58947 63662: contig of 4716 bp in length
63663 63762: gap of unknown length
63763 69752: contig of 5990 bp in length
69753 69852: gap of unknown length
69853 75209: contig of 5357 bp in length
75210 75309: gap of unknown length
75310 80744: contig of 5435 bp in length
80745 80844: gap of unknown length
80845 87050: contig of 6206 bp in length
87051 87150: gap of unknown length
87151 92697: contig of 5547 bp in length
92698 92797: gap of unknown length
92799 96711: contig of 3914 bp in length
96712 96812: gap of unknown length
103572: contig of 6761 bp in length

```

* 103573 103672: gap of unknown length
* 103673 111357: contig of 7685 bp in length
* 111358 111457: gap of unknown length
* 111458 119120: contig of 7663 bp in length
* 119121 119220: gap of unknown length
* 119221 126481: contig of 7261 bp in length
* 126482 126581: gap of unknown length
* 126582 138776: contig of 12195 bp in length
* 138777 138876: gap of unknown length
* 138877 154763: contig of 15887 bp in length
* 154764 154863: gap of unknown length
* 154864 166124: contig of 11261 bp in length
* 166125 167108: gap of unknown length
* 167109 167108: contig of 884 bp in length.

```

FEATURES

```

source
1. 167108
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="19"
   /clone="RP11-317B13"
1. 1540
   /note="assembly_name:Contig23"
misc_feature
1641. 2783
   /note="assembly_name:Contig24"
misc_feature
2884. 4448
   /note="assembly_name:Contig25"
misc_feature
4549. 6251
   /note="assembly_name:Contig26"
misc_feature
6352. 9081
   /note="assembly_name:Contig27"
misc_feature
9182. 10460
   /note="assembly_name:Contig28"
misc_feature
10561. 12863
   /note="assembly_name:Contig29"
misc_feature
12964. 15536
   /note="assembly_name:Contig30"
misc_feature
15637. 18260
   /note="assembly_name:Contig31"
misc_feature
18361. 20681
   /note="assembly_name:Contig32"
misc_feature
20782. 24034
   /note="assembly_name:Contig33"
misc_feature
24135. 27581
   /note="assembly_name:Contig34"
misc_feature
27682. 30154
   /note="assembly_name:Contig35"
misc_feature
30255. 33662
   /note="assembly_name:Contig36"
misc_feature
33763. 35826
   /note="assembly_name:Contig37"
misc_feature
35927. 339520
   /note="assembly_name:Contig38"
misc_feature
39621. 42496
   /note="assembly_name:Contig39"
misc_feature
42597. 45290
   /note="assembly_name:Contig40"
misc_feature
45391. 50042
   /note="assembly_name:Contig41"
misc_feature
50143. 54382
   /note="assembly_name:Contig42
clone end:SP6
vector_side:left"
54483. 558846
   /note="assembly_name:Contig43"
misc_feature
58947. 63662
   /note="assembly_name:Contig44"
misc_feature
63763. 69752
   /note="assembly_name:Contig45"
misc_feature
69853. 75209
   /note="assembly_name:Contig46"
misc_feature
75310. 80744
   /note="assembly_name:Contig47"

```

```

misc_feature 80845. 87050
   /note="assembly_name:Contig48"
misc_feature 87151. 92697
   /note="assembly_name:Contig49"
misc_feature 92798. 96711
   /note="assembly_name:Contig50"
misc_feature 96812. 103572
   /note="assembly_name:Contig51"
misc_feature 103673. 111357
   /note="assembly_name:Contig52"
misc_feature 111458. 119120
   /note="assembly_name:Contig53"
misc_feature 119221. 126481
   /note="assembly_name:Contig54"
misc_feature 126582. 138776
   /note="assembly_name:Contig55"
misc_feature 138877. 154763
   /note="assembly_name:Contig56"
misc_feature 154864. 166124
   /note="assembly_name:Contig57"

```

Query Match 7.6%; Score 144; DB 2; Length 167108;
 Best Local Similarity 100.0%; Pred. No. 3; 9e-64;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 569 CAGGCTCCCCCAACCAACGCGGCGAGTCGCGGCGCGCCAGTGTGTSCCA 628
    |||||
Db 23027 CAGGCTCCCCCAACCAACGCGGCGAGTCGCGGCGCGCCAGTGTGTSCCA 22368
    |||||
QY 629 GACGCGCGCGAGAGGGGCGCCAGCCGACCCAGCAGCAGTCAAGACCCCGCAGACTT 688
    |||||
Db 22967 GACGCGCGCGAGAGGGGCGCCAGCCGACCCAGCAGCAGTCAAGACCCCGCAGACTT 22308
    |||||
QY 689 CCTGTGAAGCACAAGTACCTGTG 712
    |||||
Db 22907 CCTGTGAAGCACAAGTACCTGTG 22884
    |||||

```

RESULT 29
 G31114/c 250 bp DNA linear STS 01-JAN-1997
 LOCUS human STS SHGC-19696, sequence tagged site.
 DEFINITION G31114
 ACCESSION G31114.1 GI:1758515
 VERSION G31114.1
 KEYWORDS STS; STS sequence; primer; sequence tagged site.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 250)
 AUTHORS Myers,R.M.
 JOURNAL Unpublished (1996)
 COMMENT

Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, M-344, Stanford, CA 94305, USA
 Tel: 415/7259687
 Fax: 415/7259689
 Email: myers@shgc.stanford.edu

Primer A: TCCTTTCCAGAGGTGG
 Primer B: TGCTGTCCAGAGCTGAAC
 STS size: 141
 PCR Profile:

Initial incubation: 94 degrees C for 90 seconds
 Denaturation: 94 degrees C for 15 seconds
 Annealing: 62 degrees C for 23 seconds
 Polymerization: 72 degrees C for 30 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9600
 Protocol:

* 50812 50911: gap of unknown length
* 50912 52640: contig of 1729 bp in length
* 52641 52740: gap of unknown length
* 52741 53952: contig of 1212 bp in length
* 53953 54053: gap of unknown length
* 54053 58018: contig of 3966 bp in length
* 58019 58118: gap of unknown length
* 58119 61436: contig of 3318 bp in length
* 61437 61536: gap of unknown length
* 61537 64032: contig of 2496 bp in length
* 64033 64132: gap of unknown length
* 64133 66752: contig of 2620 bp in length
* 66753 66852: gap of unknown length
* 66853 70097: contig of 3245 bp in length
* 70098 70197: gap of unknown length
* 70198 73503: contig of 3306 bp in length
* 73504 73603: gap of unknown length
* 73604 80346: contig of 6743 bp in length
* 80347 80446: gap of unknown length
* 80447 86466: contig of 6020 bp in length
* 86467 92278: gap of unknown length
* 92279 92378: contig of 5712 bp in length
* 92379 98874: gap of unknown length
* 98875 98974: contig of 6496 bp in length
* 98975 106050: contig of 7076 bp in length
* 106051 106150: gap of unknown length
* 106151 113625: contig of 7475 bp in length
* 113626 113725: gap of unknown length
* 113726 129075: contig of 15350 bp in length
* 129076 129175: gap of unknown length
* 129176 141800: contig of 12625 bp in length
* 141801 141900: gap of unknown length
* 141901 155581: contig of 13681 bp in length
* 155582 155681: gap of unknown length
* 155682 170324: contig of 14643 bp in length
* 170325 170424: gap of unknown length
* 170425 184681: contig of 14257 bp in length
* 184682 184781: gap of unknown length
* 184782 206159: contig of 21378 bp in length
* 206160 226909: gap of unknown length
* 226909: contig of 20650 bp in length.

FEATURES

source
1. .226909
Location/Qualifiers
/organism="Mus musculus"
/db_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-190A2"
/clone_lib="RPC1 mouse BAC library 23"

ORIGIN

Query Match 3.3%; Score 62; DB 2; Length 226909;
Best Local Similarity 100.0%; Pred. No. 8.1e-21;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTAATATAGAGAGGTCGACAGAAAGATCCAAAAGTGGCTGCGCGCGTCCGCC 60
Db 179014 AGCGCTAATATAGAGAGGTCGACAGAAAGATCCAAAAGTGGCTGCGCGCGTCCGCC 179073

QY 61 AG 62

Db 179074 AG 179075

RESULT 31
AF457576 993 bp mRNA linear ROD 10-APR-2002
LOCUS AF457576
DEFINITION Mus musculus death effector domain-containing protein FLAME-3
(Flame-3) mRNA, complete cds.
ACCESSION AF457576
VERSION AF457576.1 GI:20126795
KEYWORDS Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 993)
AUTHORS Zhan, Y., Hegde, R., Srinivasula, S.M., Fernandes-Alnemri, T. and Alnemri, E.S.
TITLE Death effector domain-containing proteins DEDD and FLAME-3 form nuclear complexes with the TFIIC102 subunit of human transcription factor IIIC
JOURNAL Cell Death Differ. 9 (4), 439-447 (2002)
MEDLINE 21961615
PUBMED 11965497
REFERENCE 2 (bases 1 to 993)
AUTHORS Alnemri, E.S.
TITLE Direct Submission
JOURNAL Submitted (11-DEC-2001) Microbiology and Immunology, Thomas Jefferson University, 233 S. 10th Street, Philadelphia, PA 19107, USA

FEATURES

source
1. .993
Location/Qualifiers
/organism="Mus musculus"
/db_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="f1p"
1. .993
/gene="Flame-3"
1. .993
/gene="Flame-3"
/codon_start=1
/product="death effector domain-containing protein FLAME-3"
/protein_id="P44108.1"
/db_xref="GI:20126796"

CDS

translation="MALSGSTPAQSWBEBDECLDYGMLSHMFVVGQITCELE
LAPLDEAPGAPGLARSGLELLELRGGCDSSNRLSLQLRVLRHDLPL
ARKRRVPEPERSYSPSSSKRTEDSGRRRQASSSPSSPOSDTGPPTKQRR
SGRRPSSGARORRAGLAASQHQHQLRPSSEKVTCDILRYAEVCEHGPAL
OGVASSRPPALRDLDFQATPVARSRDLSGVVCDIKSELSLAFMGDYSGLL
QALRGVPLREALRANVRALVLSVDEADYAGRRLLIMEBGRRTKAS"

ORIGIN

Query Match 3.0%; Score 56; DB 10; Length 993;
Best Local Similarity 100.0%; Pred. No. 2.1e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 TCCTATCTGAGCCCTTCTGGGCGACTACTAGTGGCGCCCTCTGTGACGCCCT 947
Db 781 TCCTATCTGAGCCCTTCTGGGCGACTACTAGTGGCGCCCTCTGTGACGCCCT 836

RESULT 32
AF543541 993 bp mRNA linear ROD 02-OCT-2002
LOCUS AF543541
DEFINITION Mus musculus death effector domain-containing DNA-binding protein 2
(Dedd2) mRNA, complete cds.
ACCESSION AF543541
VERSION AF543541.1 GI:23452816
KEYWORDS

SOURCE

Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 993)
AUTHORS Alciivar, A., Hu, S. and Yang, X.
TITLE DEDD2 is a DED-containing protein targeted to the nucleolus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 993)
AUTHORS Alciivar, A., Hu, S. and Yang, X.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2002) Abramson Family Cancer Research Institute, Univ of Pennsylvania, 421 Curie Blvd. 635 BRB1/III, Philadelphia, PA 19104-6160, USA

FEATURES	location/Qualifiers
source	1..993
gene	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	1..993
	/gene="Ded42"
	1..993
	/gene="Ded42"
	/codon_start=1
	/product="death effector domain-containing DNA-binding protein 2"
	/protein_id="AA033179.1"
	/db_xref="GI:23452817"
	/translation="MALSGRTPASWSEDELDYGMLSHMEFYVGGQITSELEL LAFILDPAGPAGGIARSGLELLLEIEREGQDSESNLRLLSOLLVLAHDLPHL ARKRRPVSPEPVYGNPSSSKRTIESCRRRPAASSSSPQOMDTPSPTRQR SRGSPSGARQRRPAGLAQQOHOHELGPSEGGKVTCDIRLRVAEYCEHPALE QGVSRPRLARLDVFGQATVLRSDIGSVVCDIKFSELSYLDAPWGDYSGAL QLARGVLTETLRVAAGREARLLVSDADYENGRRLLMEEGGRRGRAS"
ORIGIN	
Query Match	3.0%; Score 56; DB 10; Length 993;
Best Local Similarity	100.0%; Pred. No. 2.1e-17;
Matches	56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	781 TCCATATCGACGCCCTTCTGGGGGACACTACTGACGTGGCGCCCTGCTGCAGCCCT 947
OY	892 TCCATATCGACGCCCTTCTGGGGGACACTACTGACGTGGCGCCCTGCTGCAGCCCT 947
LOCOS	BC037043
RESULT 33	
LOCOS	BC037043
DEFINITION	Mus musculus death effector domain-containing DNA binding protein
ACCESSION	2, mRNA (cDNA clone IMAGE:4526883), complete cds.
VERSION	BC037043
KEYWORDS	BC037043.1 GI:22477602
ORGANISM	Mus musculus (house mouse)
SOURCE	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus. 1 (bases 1 to 1540)
AUTHORS	Krausberg, R.L., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshitsugu, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wotley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Vitaliano, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
TITLE	12477932
JOURNAL	2 (bases 1 to 1540)
REFERENCE	Strausberg, R.
AUTHORS	Direct Submission
TITLE	Submitted (23-AUG-2002) National Institutes of Health, Mammalian
JOURNAL	Gene Collection (MGC), Cancer Genomics Office, National Cancer
USA	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CNSA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nyc.mgc@nhgri.nih.gov
Ahner,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakeley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laic,P., Legaspi,R.,
Maduro,Q.L., Masello,C., Maekari,B., Mastrian,S.D., McCloskey,J.C.,
McDowall,J.J., Pearson,R., Stantiripop,S., Thomas,P.J., Touchman,J.W.,
Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 81 Row: c Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23956157
This clone has the following problem: The cds is short compared to the longest cds in the locus.

FEATURES
SOURCE

1..1540
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cdname="IMAGE:4526883"
/tissue_type="Eye, retina, mouse strain C57BL\6"
/clone_id="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1..1540
/gene="Dedd2"
/note="synonym: FLAME-3"
/db_xref="Locustid:67373"
/db_xref="WGI:1914629"
281..754
/gene="Dedd2"
/codon_start=1
/product="Dedd2 protein"
/protein_id="AAH37043.1"
/db_xref="GI:22477603"
/db_xref="Locustid:67379"
/db_xref="WGI:1914629"
/translation="MGKQAGRGVECTCPFPARAMQSWDSIDRLRVRAEYCEHGPAALL
EOGVAARRPOLARLDVFQCATAVASRDLGSVCDIKSELSYLDAFWGDYISGLAL
LOALRGVFLTEALREAVGRFAVRLLVSVDADYEAGRRIILLMEEGGRGTREAS"

CDS

ORIGIN

Query Match 3.0%; Score 56; DB 10; Length 1540;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 892 TCCATATCGACGGCTTCGTGGGGGACTACTGAGTGAGGCCCGCTGACGAGCCCT 947
|||||
Db 542 TCCATATCGACGGCTTCGTGGGGGACTACTGAGTGAGGCCCGCTGACGAGCCCT 597
|||||

RESULT 34
LOCUS AC120393/c
DEFINITION Mus musculus chromosome 7 clone RP24-312B12 map 7, *** SEQUENCING
IN PROGRESS ***, 19 unordered pieces.
AC120393
AC120393.4 GI:45736765
HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
Mus musculus (house mouse)

ORGANISM

REFERENCE

Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 190669)
Birren, B., Nusbaum, C. and Lander, E.
Mus musculus chromosome 7, clone RP24-312B12
Unpublished
2 (bases 1 to 190669)

TITLE
JOURNAL
AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Bouckgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Garde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagoe, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Landers, T., Lenoczky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mhoya, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
AUTHORS

Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 190669)

TITLE
JOURNAL
COMMENT

Submitted (05-JUL-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 25, 2004 this sequence version replaced gr:30017810.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://fcp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@road.mit.edu

Project Information
Center project name: L25953

Center clone name: 312_B_12

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 23038: contig of 23038 bp in length
* 23039 23138: gap of 100 bp
* 23138 25827: contig of 2689 bp in length
* 25827 25927: gap of 100 bp
* 25928 28740: contig of 2813 bp in length
* 28741 28840: gap of 100 bp
* 28841 45485: contig of 16645 bp in length
* 45486 45585: gap of 100 bp
* 45586 48996: contig of 3411 bp in length
* 48997 49096: gap of 100 bp
* 49097 52311: contig of 3215 bp in length
* 52312 52411: gap of 100 bp
* 52412 55723: contig of 3312 bp in length
* 55724 55823: gap of 100 bp
* 55824 69848: contig of 14025 bp in length
* 69849 69948: gap of 100 bp
* 69949 83553: contig of 13605 bp in length
* 83554 83653: gap of 100 bp
* 83654 96037: contig of 12384 bp in length
* 96038 96137: gap of 100 bp
* 96138 101113: contig of 4976 bp in length
* 101114 101213: gap of 100 bp
* 101214 103726: contig of 2513 bp in length
* 103727 103826: gap of 100 bp
* 103827 106119: contig of 2293 bp in length
* 106120 106219: gap of 100 bp
* 106220 116675: contig of 10456 bp in length
* 116676 116775: gap of 100 bp
* 116776 127768: contig of 10993 bp in length
* 127769 127868: gap of 100 bp
* 127869 135609: contig of 7741 bp in length
* 135610 135709: gap of 100 bp
* 135710 157455: contig of 21746 bp in length
* 157456 157555: gap of 100 bp
* 157556 162943: contig of 5388 bp in length
* 162944 163043: gap of 100 bp
* 163044 190669: contig of 27626 bp in length.

FEATURES
source

1. 190669
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/map="7"
/clone="RP24-312B12"
/clone_1kb="RPCT-24 Male Mouse BAC"

ORIGIN

Query Match 3.0%; Score 56; DB 2; Length 190669;
Best Local Similarity 100.0%; Pred. No. 1,2e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 892 TCTATCTGAGCGCCCTTCTGGGCGACTACCTGAGTGGCGCCCTGCTGAGCGCCCT 947
Db 177706 TCTATCTGAGCGCCCTTCTGGGCGACTACCTGAGTGGCGCCCTGCTGAGCGCCCT 177651

RESULT 35

AC079490/c 226909 bp DNA linear HTG 02-SEP-2000
LOCUS Mus musculus clone RP23-190A2, WORKING DRAFT SEQUENCE, 53 unordered
DEFINITION pieces.
AC079490
AC079490.1 GI:9964855
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 226909)
DOE Joint Genome Institute.
Sequencing of Mouse
Unpublished
2 (bases 1 to 226909)
DOE Joint Genome Institute.
Direct Submission
Submitted (02-SEP-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----genome center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1809834
Center Clone name: RPCT-23_190A2

Summary Statistics
Consensus quality: 180161 bases at least Q40
Consensus quality: 198601 bases at least Q30
Consensus quality: 203731 bases at least Q20
Estimated insert size: 209000; agarose-fp estimation
Estimated insert size: 221709; sum-of-contigs estimation
Quality coverage: 6.2 in Q20 bases; agarose-fp estimation
Quality coverage: 5.84 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1121 1120: contig of 1120 bp in length
1220 1220: gap of unknown length
1221 1220: contig of 1354 bp in length
2575 2574: gap of unknown length
2674 2674: contig of 1388 bp in length
4063 4062: contig of 1388 bp in length
4163 4162: gap of unknown length
5667 5666: contig of 1504 bp in length
5767 5766: gap of unknown length
6895 6895: contig of 1129 bp in length
6896 6895: gap of unknown length
6995 6995: contig of 1030 bp in length
8025 8025: gap of unknown length
8125 8125: contig of 1263 bp in length
9388 9388: gap of unknown length
9489 9488: gap of unknown length
10710 10710: contig of 1222 bp in length
10810 10810: gap of unknown length
10811 10811: contig of 1001 bp in length
11812 11811: gap of unknown length
11911 11911: gap of unknown length
12965 12965: contig of 1054 bp in length
13065 13065: gap of unknown length
14268 14268: contig of 1203 bp in length
14369 14368: gap of unknown length
14369 14368: contig of 1478 bp in length
15847 15846: gap of unknown length
15946 15946: contig of 1273 bp in length
17219 17219: gap of unknown length
17319 17319: gap of unknown length
18672 18672: contig of 1353 bp in length
18773 18772: gap of unknown length
19999 19999: contig of 1227 bp in length
20099 20099: gap of unknown length
20100 20100: contig of 1360 bp in length
21460 21459: gap of unknown length
21559 21559: gap of unknown length
22622 22622: contig of 1063 bp in length
22722 22722: gap of unknown length
23726 23726: contig of 1004 bp in length
23826 23826: gap of unknown length
25149 25149: contig of 1323 bp in length

25150 25249: gap of unknown length
25250 26839: contig of 1590 bp in length
26840 26839: gap of unknown length
26940 26939: contig of 1691 bp in length
28631 28730: gap of unknown length
28731 30265: contig of 1535 bp in length
30266 30365: gap of unknown length
30366 31703: contig of 1338 bp in length
31704 31803: gap of unknown length
31804 32952: contig of 1148 bp in length
32952 33051: gap of unknown length
33052 33299: contig of 2278 bp in length
33299 35330: gap of unknown length
35330 35429: gap of unknown length
35429 38496: contig of 3067 bp in length
38496 38596: gap of unknown length
38597 40521: contig of 1925 bp in length
40521 40622: gap of unknown length
40622 42212: contig of 1591 bp in length
42212 42312: gap of unknown length
42312 44441: contig of 2129 bp in length
44441 44541: gap of unknown length
44542 46397: contig of 1856 bp in length
46397 46497: gap of unknown length
46498 46585: contig of 2168 bp in length
46586 48785: gap of unknown length
48786 50811: contig of 2026 bp in length
50812 50911: gap of unknown length
50912 52640: contig of 1739 bp in length
52641 52740: gap of unknown length
52741 53952: contig of 1212 bp in length
53953 54052: gap of unknown length
54053 58018: contig of 3966 bp in length
58019 58119: gap of unknown length
58119 61436: contig of 3318 bp in length
61437 61536: gap of unknown length
61537 64032: contig of 2496 bp in length
64033 64132: gap of unknown length
64133 66752: contig of 2620 bp in length
66753 66852: gap of unknown length
66853 70097: contig of 3245 bp in length
70098 70197: gap of unknown length
70198 73503: contig of 3306 bp in length
73504 73603: gap of unknown length
73604 80346: contig of 6743 bp in length
80347 80446: gap of unknown length
80447 86466: contig of 6020 bp in length
86467 86566: gap of unknown length
86567 92278: contig of 5712 bp in length
92279 92378: gap of unknown length
92379 98874: contig of 6436 bp in length
98875 98974: gap of unknown length
98975 106050: contig of 7076 bp in length
106051 106150: gap of unknown length
106151 113625: contig of 7475 bp in length
113626 113725: gap of unknown length
113726 129075: contig of 15350 bp in length
129076 129175: gap of unknown length
129176 141800: contig of 12625 bp in length
141801 141900: gap of unknown length
141901 155581: contig of 13681 bp in length
155582 155681: gap of unknown length
155682 170324: contig of 14643 bp in length
170325 170424: gap of unknown length
170425 184681: contig of 14257 bp in length
184682 184781: gap of unknown length
184782 206159: contig of 21378 bp in length
206160 206259: gap of unknown length
226909 226909: contig of 20650 bp in length.

1. 226909
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

FEATURES
source

```

/clone="RP23-190A2"
/clone_1lb="RPC1 mouse BAC library 23"
ORIGIN
Query Match      2.2%; Score 41; DB 2; Length 226909;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 AGGCGTATATATAGAGAGGCGCAGAGAAATCCAAAACAA 41
Db      12604 AGGCGTATATATAGAGAGGTCGAGAGATCCAAAACAA 12564

RESULT 36
AR339362/c
LOCUS      AR339362      1067 bp      DNA      linear      PAT 17-AUG-2003
DEFINITION Sequence 853 from patent US 6569662.
ACCESSION AR339362
VERSION    AR339362.1 GI:33726219
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown:
            Unclassified.
REFERENCE  1 (bases 1 to 1067)
AUTHORS    Tang,Y.T., Zhou,P. and Drmanac,R.T.
TITLE      Nucleic acids and polypeptides
JOURNAL    Patent: US 6569662-A 853 27-MAY-2003;
FEATURES   Location/Qualifiers
            1..1067
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match      2.0%; Score 37; DB 6; Length 1067;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1760 ACATGGGGCCAAACAGTAACAGCCACTTCTTG 1796
Db      979 ACATGGGGCCAAACAGTAACAGCCACTTCTTG 943

RESULT 37
AC140075
LOCUS      AC140075      197751 bp      DNA      linear      HTG 27-MAR-2003
DEFINITION Rattus norvegicus clone CH230-370D22, WORKING DRAFT SEQUENCE, 29
            unordered pieces.
ACCESSION AC140075
VERSION    AC140075.4 GI:28894562
KEYWORDS   HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 197751)
            Muzny,D.,Marle,M., Metzker,M.,Lee,S., Abramson,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alibrooks,S., Amin,A., Angiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biewald,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacho,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebreyesgizb,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,

```

```

Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A.B., Hogues,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensbush,L., Louisedge,H., Lozada,R.J., Lu,X., Ma,J.,
Manshwar,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M., McNeill,T., Meenen,E., Milosavljevic,A.,
Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K.,
Morris,S., Mundiadae,M., Murphy,M., Nair,L., Nankervis,C., Neal,D.,
Newton,N., Nguyen,N., Norris,S., Nwaokemehe,O., Okunnu,G.,
Oiarumpasagoon,A., Pal,S., Parke,K., Pasternak,S., Paul,H.,
Perez,A., Perez,L., Pfankoch,C., Plopper,F., Poindexter,A.,
Popovich,D., Primus,E., Pu,L.-L., Puzo,M., Quiroz,J., Rachlin,E.,
Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y.,
Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A.,
Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S.,
Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A.,
Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E.,
Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A.,
Syatek,A., Tabot,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S.,
Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseña,D.,
Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J.,
Warren,R., Wei,X., White,P., Williams,G., Willson,R., Wlecyk,R.,
Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S.,
Yan,U., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X.,
Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R.,
Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Unpublished
Direct Submission
2 (bases 1 to 197751)
Worley,K.C.
Submitted (21-FEB-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 197751)
Worley,K.C.
Direct Submission
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 10, 2003 this sequence version replaced gi:28475367.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KDWU
Center clone name: CH230-370D22
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 188105 bases at least Q40
Consensus quality: 192608 bases at least Q30
Consensus quality: 196096 bases at least Q20
Estimated insert size: 184810; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```


REFERENCE 3 (bases 1 to 223642)
 AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 9, 2002 this sequence version replaced gi:23267817.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GOVZ
 Center clone name: CH230-201C19

----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 198355 bases at least Q40
 Consensus quality: 200628 bases at least Q30
 Consensus quality: 202141 bases at least Q20
 Estimated insert size: 200365; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 17684: contig of 17684 bp in length
 * 17685 17784: gap of unknown length
 * 17785 109488: contig of 91704 bp in length
 * 109489 109588: gap of unknown length
 * 109589 218503: contig of 108915 bp in length
 * 218504 218603: gap of unknown length
 * 218604 219707: contig of 1104 bp in length
 * 219708 219807: gap of unknown length
 * 219808 223642: contig of 3835 bp in length.
 Location/Qualifiers

FEATURES
 source
 1..223642
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-201C19"
 2773..3305
 /note="clone_boundary"

misc_feature
 clone_end:17
 site:ECORI
 end_sequence:BH330265"

misc_feature
 17785..18888
 /note="wgs_contig"

misc_feature
 135936..138590
 /note="wgs_contig"

ORIGIN

Query Match 1.3%; Score 24; DB 2; Length 223642;
 Best Local Similarity 100.0%; Pred. No. 0.97;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1239 TCCTCCAGACCTCTCTACCCC 1262
 Db 21381 TCCTCCAGACCTCTCTACCCC 21358

RESULT 39
 BC075935
 LOCUS
 DEFINITION Danio rerio zgc:92202, mRNA (cDNA clone MGC:92202 IMAGE:7051835), complete cds.
 ACCESSION BC075935
 VERSION BC075935.1 GI:49904383
 KEYWORDS MGC.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Cassavani, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1821)

JOURNAL
 PUBLISHED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Direct Submission
 Submitted (06-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapds-r@mail.nih.gov
 Tissue Procurement: Len Zon, Harvard
 CDNA Library Preparation: Open BioSystems
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.hgsc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>
 Series: IRAC Plate: 179 Row: 1 Column: 16
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES
 source
 1..1821
 /organism="Danio rerio"
 /mol_type="mRNA"

```

gene
  /db_xref="taxon:7955"
  /clone="MGC:92202 IMAGE:7051835"
  /issue_type="Whole, adult zebrafish"
  /clone_lib="NIH ZGC_7"
  /lab_host="DH10B"
  /note="Vector: pExpress1"
  1..1821
  /gene="zgc:92202"
  /db_xref="LocustID:436912"
  268..1482
  /gene="zgc:92202"
  /product="zgc:92202"
  /protein_id="AAH75935.1"
  /db_xref="GI:49904384"
  /db_xref="LocustID:436912"
  /translation="MTSQQHGNANPALISLSPONSSSGOGRHISRPIDSYRSRGLSSP
  RGRWVAASSSGTTHGMSLALSRLAPASSNSSSTRPAPSGVGEWPEAPDADYGL
  YSHRPFVDIVGAOLTHRDVAVLSFLVDVIDEYRGGISGRDFILALEROGRCDETN
  FRHVLQLRIITRHDLIPYTLAKRQIVCEPDVDKLESTVAVYPSRGTADAQQGTP
  HRRHTQPLICPPSGPOVCPPRAPKAPPPSPKRRKSHTTADCRKQTCIDRLRARA
  EYGHESALQGNVPSNKOALEROFPERNQANTILKSRDLGSIICDKFSSELYLDAP
  MRYINGSLLEALKGVFITDSLKQAVGHEAIKLVVDEEDYQAGRKILRNLVAGCA
  GACTGSRBEPPLS"

ORIGIN
  Query Match      1.2%; Score 23; DB 5; Length 1821;
  Best Local Similarity 100.0%; Pred. No. 5.4;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      868 TGTGACATCAAGTTCTCAGAGCT 890
Db      1228 TGTGACATCAAGTTCTCAGAGCT 1250

RESULT 40
AC091788      92932 bp      DNA      linear      PRI 09-MAR-2002
LOCUS      Homo sapiens BAC clone RP11-400A23 from 2, complete sequence.
DEFINITION
AC091788
AC091788.6      GI:19310343
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
Homo sapiens
ORGANISM
REFERENCE
AUTHORS
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
JOURNAL
Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE
99063792
PUBMED
9847074
REFERENCE
AUTHORS
Trani,L. and Corton,M.
The sequence of Homo sapiens BAC clone RP11-400A23
JOURNAL
Unpublished (2001)
REFERENCE
AUTHORS
Waterston,R.H.
Direct Submission
Submitted (06-JUN-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 92932)
Waterston,R.H.
Direct Submission
Submitted (21-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 92932)
Waterston,R.
Direct Submission
Submitted (09-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

```

```

COMMENT
On Mar 9, 2002 this sequence version replaced gi:18855146.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplen@watsn.wustl.edu
-----
Summary Statistics
Center project name: H_NH0400A23
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCT-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Feigenz.B.,
Tateno,M., Cattanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-9F10, 2000 bp overlap; the
clone sequenced to the right is RP11-19A8, 2000 bp overlap.

Single plasmid region exists between 89575 and 89677. Polymorphisms
exist between AC087816 and AC091788.
Location/Qualifiers
1..92932
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-400A23"
/clone_lib="RPCT-11"
1..1053
/rpc_family="L1"
repeat_region
2066..2194
/rpc_family="L2"
repeat_region
2736..3030
/rpc_family="L1"
repeat_region
3572..3865
/rpc_family="Alu"
3896..3936
/rpc_family="MALR"
3940..4181
/rpc_family="L1"
repeat_region
4203..4648
/rpc_family="ERV1"
4788..4878
/rpc_family="(TTAA)n"
5492..55620
/rpc_family="MALR"
repeat_region
6008..6035

```



```

repeat_region      /rpt_family="AT_rich"
                    6120. .6159
repeat_region      /rpt_family="(CA)n"
                    6577. .6788
repeat_region      /rpt_family="CRI"
                    6910. .6965
repeat_region      /rpt_family="MIR"
                    7324. .7549
repeat_region      /rpt_family="MERL_type"
                    7583. .7709
repeat_region      /rpt_family="(TTCC)n"
                    7731. .7753
repeat_region      /rpt_family="L1"
                    7754. .8087
repeat_region      /rpt_family="L1"
                    9813. .9971
repeat_region      /rpt_family="Alu"
                    10162. .10464
repeat_region      /rpt_family="Alu"
                    11489. .11561
repeat_region      /rpt_family="MIR"
                    11563. .11673
repeat_region      /rpt_family="CRI"
                    11667. .11711
repeat_region      /rpt_family="MIR"
                    12629. .12724
repeat_region      /rpt_family="L2"
                    13105. .13603
repeat_region      /rpt_family="L1"
                    14129. .14155
repeat_region      /rpt_family="(TG)n"
                    14582. .14665
repeat_region      /rpt_family="MIR"
                    14789. .14899
repeat_region      /rpt_family="MERL_type"
                    15090. .15137
repeat_region      /rpt_family="MIR"
                    15363. .16129
repeat_region      /rpt_family="L1"
                    16130. .16425
repeat_region      /rpt_family="Alu"
                    16426. .16436
repeat_region      /rpt_family="L1"
                    16437. .16745
repeat_region      /rpt_family="Alu"
                    16746. .17016
repeat_region      /rpt_family="L1"
                    17022. .17179
repeat_region      /rpt_family="Alu"
                    17183. .17474
repeat_region      /rpt_family="L1"
                    17484. .17616
repeat_region      /rpt_family="L1"
                    18559. .18687
repeat_region      /rpt_family="MIR"
                    19885. .19976
repeat_region      /rpt_family="L1"
                    20140. .20516
repeat_region      /rpt_family="BRV1"
                    20637. .20754
repeat_region      /rpt_family="MIR"
                    20912. .21170
repeat_region      /rpt_family="L2"
                    21184. .21581
repeat_region      /rpt_family="L2"
                    21747. .21954
repeat_region      /rpt_family="MIR"
                    22398. .23141
repeat_region      /rpt_family="MIR"
                    24551. .24575
repeat_region      /rpt_family="(TC)n"
                    25072. .25186
repeat_region      /rpt_family="CT-rich"

```

```

repeat_region      26654. .26688
                    /rpt_family="GA-rich"
repeat_region      26805. .26941
                    /rpt_family="L1"
repeat_region      26942. .27271
                    /rpt_family="Alu"
repeat_region      27272. .27670
                    /rpt_family="L1"
repeat_region      27689. .27950
                    /rpt_family="L1"
repeat_region      27951. .28470
                    /rpt_family="MIR"
repeat_region      28471. .29529
                    /rpt_family="L1"
repeat_region      29537. .29880
                    /rpt_family="L2"
repeat_region      30143. .30264
                    /rpt_family="CRI"

```

```

Query Match      1.2%; Score 23; DB 9; Length 92932;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1221 ACCATTTCAGCTGAGTTTCCTT 1243
Db      79796 ACCATTTCAGCTGAGTTTCCTT 79818

```

```

RESULT 41
BX248385/c      125523 bp      DNA      linear      VRT 19-NOV-2003
LOCUS      Zebrafish DNA sequence from clone CH211-229N2 in linkage group 8,
DEFINITION      complete sequence.
ACCESSION      BX248385
VERSION      BX248385.6 GI:38453664
KEYWORDS      HTG.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
      Cypriniformes; Cyprinidae; Danio.
      1 (bases 1 to 125523)
REFERENCE      Submitted (19-NOV-2003) Wellcome Trust Sanger Institute, Hinxton,
AUTHORS      Cambridgehire, CB10 1SA, UK. E-mail enquiries:
TITLE      zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
JOURNAL      On Nov 19, 2003 this sequence version replaced gi:38304102.
COMMENT      -----
      Center: Wellcome Trust Sanger Institute
      Center code: SC
      Web site: http://www.sanger.ac.uk
      Contact: zfish-help@sanger.ac.uk
      -----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived

Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhirong Bao and Sean Eddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rexio/fishmask.shtml CH211-229N2 is from a CHORI-211 BAC library

FEATURES
source

1.125523
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-229N2"
/clone_1ib="CHORI-211"

ORIGIN

Query Match 1.2%; Score 23; DB 5; Length 125523;

Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 868 TGTGACATCAAGTTCTCAGAGCT 890
|||||
Db 63077 TGTGACATCAAGTTCTCAGAGCT 63055

RESULT 42

LOCUS CQ411697 838 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 18768 from Patent WO0170979.
ACCESSION CQ411697
VERSION CQ411697.1 GI:41319478
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Lee,J. and Lillie,J.
TITLE Genes, compositions, kits, and method for identification, assessment, prevention, and therapy of ovarian cancer
JOURNAL Patent: WO 0170979-A 18768 27-SEP-2001,
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source

1.838
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.2%; Score 22; DB 6; Length 838;

Best Local Similarity 100.0%; Pred. No. 20;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 GGGAGGAGATGAGTGCCTGGA 179
|||||
Db 381 GGGAGGAGATGAGTGCCTGGA 360

RESULT 43

LOCUS AX377832 1339 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 27 from Patent WO0212358.
ACCESSION AX377832
VERSION AX377832.1 GI:19573923
KEYWORDS

SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1
AUTHORS Gillen,C., Wetzeis,I., Whendt,S., Weihe,E. and Schaefer,M.K.
TITLE Screening method
JOURNAL Patent: WO 0212338-A 27 14-FEB-2002;
Gruenthal GmbH (DE)
FEATURES
source

1.1339
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"

ORIGIN

Query Match 1.2%; Score 22; DB 6; Length 1339;

Best Local Similarity 100.0%; Pred. No. 19;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1449 CTGCCCTGGCGCTTCCGAGC 1470
|||||
Db 1018 CTGCCCTGGCGCTTCCGAGC 1039

RESULT 44

LOCUS AF171096 2373 bp mRNA linear INV 02-OCT-2000
DEFINITION Strongylocentrotus purpuratus N-ethylmaleimide-sensitive factor
ACCESSION AF171096 AF038010
VERSION AF171096.1 GI:10443499
KEYWORDS

SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidae; Echinoidae; Echinacea; Echinoidae;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE 1 (bases 1 to 2373)
AUTHORS Timmers,K., Kolosova,I., Kingsley,D. and Zimmerberg,J.
TITLE Sea urchin N-ethylmaleimide-sensitive factor

JOURNAL Unpublished
AUTHORS Unpublished
TITLE Unpublished

REFERENCE 2 (bases 715 to 2373)
AUTHORS Whalley,T., Suprynowicz,F., Tahara,M., Timmers,K., Kingsley,D., Scheller,R. and Zimmerberg,J.
TITLE Sea urchin egg exocytotic cortical granules are primed for fusion at a stage that does not require NSF

JOURNAL Unpublished
AUTHORS Unpublished
TITLE Unpublished

REFERENCE 3 (bases 715 to 2373)
AUTHORS Kingsley,D., Timmers,K. and Zimmerberg,J.
TITLE Direct Submision
JOURNAL Submitted (10-DEC-1997) NICHD, NIH, Room 10D14, 10 Center Drive, Bethesda, MD 20892-1855, USA

REFERENCE 4 (bases 1 to 2373)
AUTHORS Timmers,K., Kolosova,I., Kingsley,D. and Zimmerberg,J.
TITLE Direct Submision
JOURNAL Submitted (22-JUN-1999) LCMR, NICHD, Bethesda, MD 20892-1855, USA

COMMENT On Oct 2, 2000 this sequence version replaced gi:4104670.
FEATURES
source

CDS

1.2373
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/feature_type="ovary"
88..2328
/codon_start=1
/product="N-ethylmaleimide-sensitive factor"
/protein_id="AA017479.1"
/db_xref="GI:10443500"
/translation="MSHREKAVRCPTQLSTNRAVSDKAPARIDYVOYQALPG
PSVFAVAPDLKSGEWFSLPQRKMAISIDQIQAAPYRDPSCQYISITVEVD
FLQKNTTNDAPDSMDKADFLSPDRAVPSVEQDLVCSFYNNKGLAVKVSVEIDP
AIIKGGKPAKRNADIVRIQVLTONTQVTEKAGSLANTGSKSGNSQSIINPMD
FKMGIIGIDKFSQI FRAPASRAVPEFVEQIGAHYKGIILFDPGKGLMART
IGKMLNSRPQIINGEPIANKVGSSEANIRTLFAAEDQKMGNSGHIITIPET
DAICORGSMSGSTGVHDTVVMDLSKITGVQELNNVLLIGMTRKDLIDDLALPRGR

```

ORIGIN
Query Match      1.2%; Score 22; DB 3; Length 2373;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1059 GCTGATGAGGAGGAGGAGGAGG 1080
       |||
Db      2292 GCTGATGAGGAGGAGGAGGAGG 2313

RESULT 45
LOCUS      COB42142                2718 bp    DNA        linear    PAT 02-AUG-2004
DEFINITION Sequence 789 from Patent EP1440981.
ACCESSION  COB42142
VERSION     COB42142.1    GI:50893929
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS     1
            Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
            Yamamoto,J., Isono,Y., Nagai,K. and Irie,R.
TITLE       Full-length human cdna
JOURNAL     Patent: EP 1440981-A 789 28-JUL-2004;
            Research Association for Biotechnology (JRP)
FEATURES
source
            1. 2718
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"

ORIGIN
Query Match      1.2%; Score 22; DB 6; Length 2718;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1262 CCAGGTGTGCCCCCTTAGCCTC 1283
       |||
Db      870 CCAGGTGTGCCCCCTTAGCCTC 891

RESULT 46
LOCUS      AK125175                2718 bp    mRNA        linear    PRI 09-SEP-2003
DEFINITION Homo sapiens cDNA FLJ43185 fls, clone FCBBF3021940, weakly similar
            to SYNAPDIN 1.
ACCESSION  AK125175
VERSION     AK125175.1    GI:34531177
KEYWORDS    oligo capping, fls (full insert sequence).
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS     1
            Taahiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
            Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
            Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
            Yamauchi,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
            Kanda,K., Magatsuna,M., Murakawa,K., Kanehori,K.,
            Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
            Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isogai,T.
TITLE       NEDO human cDNA sequencing project

```

```

JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 2718)
AUTHORS     Isogai,T. and Yamamoto,J.
TITLE       Direct Submission
JOURNAL     Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
            Kazuoka-Kamatari, Kisarazu, Chiba 292-0818, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
            NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.

FEATURES
source
            1. 2718
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="FCBBF3021940"
               /tissue_type="brain"
               /clone_lib="FCBBF3"
               /dev_stage="fetal"
               /note="Cloning vector: pME18SFL3"
               353..772
               /note="unnamed protein product"
               /codon_start=1
               /protein_id="BAC86071.1"
               /db_xref="GI:34531178"
               /translation="MTNPQVQASLAPARAGPPPTPRAPPDSCGHAPSGFPALG
               LRTAORPFPSSPFSASGRLRGPRVARGPAPASSPTGLPAYLPAAALDSQTSAPT
               VSPVTRPRVARKTPRVSLAGLETISLLHQQLND"

ORIGIN
Query Match      1.2%; Score 22; DB 9; Length 2718;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1262 CCAGGTGTGCCCCCTTAGCCTC 1283
       |||
Db      870 CCAGGTGTGCCCCCTTAGCCTC 891

RESULT 47
LOCUS      AL359829/c              112782 bp    DNA        linear    HTG 09-MAR-2001
DEFINITION Homo sapiens chromosome 10 clone RP13-169M5, WORKING DRAFT
            SEQUENCE, 20 unordered pieces.
ACCESSION  AL359829
VERSION     AL359829.5    GI:133274765
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS     1
            Sims,S.
TITLE       Direct Submission
JOURNAL     Submitted (06-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
            On Mar 12, 2001 this sequence version replaced gi:9856796.

COMMENT
            ----- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            ----- Project Information
            Center project name: DB159M5
            ----- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: pGAP4; 108752; 100% of reads

```

Chemistry: Dye-terminator Big Dye, 100% of reads
 Consensus quality: 101989 bases at least Q40
 Consensus quality: 106629 bases at least Q30
 Consensus quality: 109120 bases at least Q20
 Insert size: 110882; sum-of-contigs
 Insert size: 166129; 8.1% error; agarose-fp
 Quality coverage: 3.01x in Q20 bases; sum-of-contigs Quality
 coverage: 2.86x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 20 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      8878: contig of 8878 bp in length
*      8879      8978: gap of 100 bp
*      8879      15097: contig of 6119 bp in length
*      15098      15197: gap of 100 bp
*      15198      30569: contig of 15372 bp in length
*      30570      30669: gap of 100 bp
*      30670      32841: contig of 2172 bp in length
*      32842      32941: gap of 100 bp
*      32941      35240: contig of 2299 bp in length
*      35240      35341: gap of 100 bp
*      35341      49709: contig of 14369 bp in length
*      49710      49809: gap of 100 bp
*      49810      52990: contig of 3181 bp in length
*      52991      53090: gap of 100 bp
*      53091      60959: contig of 7869 bp in length
*      60960      61059: gap of 100 bp
*      61060      66156: contig of 5097 bp in length
*      66157      66256: gap of 100 bp
*      66257      69020: contig of 2764 bp in length
*      69021      69120: gap of 100 bp
*      69121      79014: contig of 9894 bp in length
*      79015      82342: contig of 3228 bp in length
*      82343      82442: gap of 100 bp
*      82443      85853: contig of 3410 bp in length
*      85853      85952: gap of 100 bp
*      85953      87993: contig of 2041 bp in length
*      87994      88093: gap of 100 bp
*      88094      90353: contig of 2260 bp in length
*      90354      90453: gap of 100 bp
*      90454      93029: contig of 2576 bp in length
*      93030      93129: gap of 100 bp
*      93130      99196: contig of 6067 bp in length
*      99197      99296: gap of 100 bp
*      99297      102316: contig of 2920 bp in length
*      102317      108820: contig of 6504 bp in length
*      108821      108920: gap of 100 bp
*      108921      112782: contig of 3862 bp in length.

```

FEATURES

source

```

1.112782
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP13-169M5"
/clone_11b="RPC1-13.1"
1.8878
/note="assembly fragment:00018
fragment_chain:1"
misc_feature
8979.15097
/note="assembly fragment:00735
fragment_chain:1"
misc_feature
15198.30569
/note="assembly fragment:01659
fragment_chain:1"

```

```

misc_feature      30670..32841
/note="assembly fragment:00525
fragment_chain:1"
misc_feature      32942..35240
/note="assembly fragment:00804
fragment_chain:2"
misc_feature      35341..49709
/note="assembly fragment:01230
fragment_chain:2"
misc_feature      49810..52990
/note="assembly fragment:01390
fragment_chain:3"
misc_feature      53091..60959
/note="assembly fragment:00339
fragment_chain:3"
misc_feature      61060..66156
/note="assembly fragment:00338"
66257..69020
/note="assembly fragment:00395"
69121..79014
/note="assembly fragment:00453"
79115..82342
/note="assembly fragment:00611"
82443..85852
/note="assembly fragment:00733"
85953..87993
/note="assembly fragment:00845"
88094..90353
/note="assembly fragment:01007"
90454..93029
/note="assembly fragment:01069"
93130..99196
/note="assembly fragment:01440"
99297..102316
/note="assembly fragment:01460"
102317..108820
/note="assembly fragment:01473"
108921..112782
/note="assembly fragment:01489"
misc_feature
misc_feature
misc_feature

```

ORIGIN

Query Match 1.2%: Score 22; DB 2; Length 112782;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1262 CCAGGTGTCCTTACGCTC 1283

Db 22006 CCAGGTGTCCTTACGCTC 21985

```

RESULT 48
AC143567/c      122619 bp      DNA      linear      HTG 09-APR-2003
LOCUS      Macaca mulatta clone CH250-267014, *** SEQUENCING IN PROGRESS ***.
ACCESSION      AC143567
VERSION      AC143567.1 GI:29648960
KEYWORDS      HTG; HTGS PHASE2; HTGS_PGI.
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM      Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
1 (bases 1 to 122619)
Caucos M. and Milosavljevic A.
pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo R. and Guefield D. (Eds.);
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2 (bases 1 to 122619)
Milosavljevic A., Sodergren E., Caucos M., Li B., Jackson A.R.,
Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C., Albrooke S.L.,

```

Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbata, J., Benton, J., Bimge, K., Blankenburg, K., Bonini, D., Bouck, J., Bowe, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Egan, A., Eshart, C., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorelli, J.H., Guvera, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homai, P., Howard, S., Huber, J., Huylk, S., Hume, J., Joshkhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, J.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louiege, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Marondei, T., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nokenko, S., Ogun, H., Okunolu, G., Otagu, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Pitman, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoochetri, N., Sison, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swatek, A., Taber, F., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telitod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Veer, V., Villalon, D., Vanson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Zuchelapati, R., Weinstein, G., and Gibbs, R.

Direct Submission
Unpublished
3 (bases 1 to 122619)
Morley, K.C.

Submitted (09-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: LBLK
Center clone name: CH250-267J14

Summary Statistics
Chemistry: Dye-terminator Big Dye: inf of reads
Consensus quality: 12732 bases at least Q40
Consensus quality: 13630 bases at least Q30
Consensus quality: 14884 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
NOTE: (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
NOTE: The contigs are based on the application
* of the PGI method using the Human genome (NCBI build 31)
* as the comparative genome.
NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 122619: contig of 122619 bp in length.
Location/Qualifiers
1. 122619
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/db_xref="taxon:9544"
/clone="CH250-267J14"
1. 122619
/note="assembly name: CH250-267J14.1B
CONFIDENCE: 0.67"

ORIGIN
Query Match 1.2%; Score 22; DB 2; Length 122619;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1166 GACCATCTACCCCTTGACAG 1167
76497 GACCATCTACCCCTTGACAG 76476

RESULT 49
AC143568/c 122976 bp DNA linear HTG 09-APR-2003
LOCUS Macaca mulatta clone CH250-267J15, *** SEQUENCING IN PROGRESS ***
AC143568
DEFINITION AC143568.1 GI:29648961
ACCESSION AC143568
VERSION HTG: HTGS PHASE2: HTGS PGI.
KEYWORDS Macaca mulatta (rhesus monkey)
SOURCE Macaca mulatta
ORGANISM Macaca mulatta (rhesus monkey)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
1 (bases 1 to 122976)
Gaurus, M. and Milosavljevic, A.
Pooled genomic indexing (PGI): mathematical analysis and experiment
design
(in) Guigo, R. and Gusfield, D. (Eds.):
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
2 (bases 1 to 122976)

REFERENCE
AUTHORS Milosavljevic, A., Sodergren, E., Gaurus, M., Li, B., Jackson, A.R.,
Adams, C., Adio-Ochola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L.,
Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbata, J.,
Benton, J., Bimge, K., Blankenburg, K., Bonini, D., Bouck, J.,
Bowe, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D.,
Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Egan, A., Eshart, C., Edwards, C.C.,
Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D.,
Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J.,
Garcia, A., Garner, T., Garza, N., Gill, R., Gorelli, J.H., Guvera, W.,
Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K.,
Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O.,
Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homai, P.,
Howard, S., Huber, J., Huylk, S., Hume, J., Joshkhes, I., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, J.C.,
Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W.,
Louiege, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapa, P., Marondei, T., Martin, R.,
Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P.,
Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G.,

Miner,Z., Mitchell,T., Nohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweto,S., Ogul,M., Okunolu,G., Oragunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Prims,E., Pu,L., R. Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shm,C., Shoshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Sytek,A., Tabor,P., Tamerias,A., Tamerias,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Zuchterlapati,R., Weinstein,G. and Gibbs,R.

Direct Submission
Unpublished
3 (bases 1 to 122976)
Worley,K.C.
Direct Submission
Submitted (09-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: LBL
Center clone name: CH250-267J15

----- Summary Statistics
Chemistry: Dye-terminator Big Dye: int% of reads
Chemistry: Dye-terminator Big Dye: int% of reads
Consensus quality: 13472 bases at least Q40
Consensus quality: 14652 bases at least Q30
Consensus quality: 16106 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: The contigs are based on the application
* of the pgi method using the Human genome (NCBI build 31)
* as the comparative genome.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 122976: contig of 122976 bp in length.
Location/Qualifiers
1.122976
/organism="Macaca mulatta"
/mol_type="genomic DNA"
/db_xref="taxon:9544"
/clone="CH250-267J15"
1.122976
/note="assembly name:CH250-267J15.1B
CONFIDENCE: 0.67"

FEATURES
source
misc_feature
1.2%; Score 22; DB 2; Length 122976;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Query Match 1.2%; Score 22; DB 2; Length 122976;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1166 GACCATCTCTACCCCTTGACAG 1187
|||||
Db 64549 GACCATCTCTACCCCTTGACAG 64528

RESULT 50
AP001976/c 141272 bp DNA linear PRI 23-JAN-2001
LOCUS
DEFINITION
Homo sapiens genomic DNA, chromosome 11q, clone:RP11-179A16.
AP001976
AP001976.4 GI:12381917
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE
1 (bases 1 to 141272)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 141,272 genomic DNA of 11q
JOURNAL
Published Only in DataBase (2000)
2 (bases 1 to 141272)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-MAY-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jan 22, 2001 this sequence version replaced gi:9927283.
STS Markers (ePCR):
22543 22694 D11S2068 11 Homo sapiens.
Location/Qualifiers
1.141272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-179A16"

FEATURES
source
1.141272
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-179A16"

ORIGIN
Query Match 1.2%; Score 22; DB 9; Length 141272;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 681 CAGACCTTCTCTGAGGCAAA 702
|||||
Db 90540 CAGACCTTCTCTGAGGCAAA 90519

Search completed: March 24, 2005, 00:19:45
Job time : 8486.52 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 16:37:44 ; Search time 6748.99 Seconds
(without alignments)
10620.110 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 1883
Sequence: 1 aggcgcatacatagagaag.....ctgaaaggcatagtg99g 1883

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapect 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

EST: *
1: gb_esc1: *
2: gb_esc2: *
3: gb_esc3: *
4: gb_esc4: *
5: gb_esc5: *
6: gb_esc6: *
7: gb_esc7: *
8: gb_esc8: *
9: gb_esc9: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1082	57.5	1866	3	CR619301 full-1eng
2	1077	57.2	1860	3	CR593642 full-1eng
3	1047	55.6	1894	3	CR625070 full-1eng
4	1002	53.2	1053	3	CR604307 full-1eng
5	731	38.8	1062	5	BM908148 full-1eng
6	731	38.8	1510	3	CR599685 full-1eng
7	731	38.8	1586	3	CR590827 full-1eng
8	708	37.6	1090	5	BM920838 full-1eng
9	680	36.1	1161	5	BM920838 full-1eng
10	629	33.4	993	4	BM551411 full-1eng
11	627	33.3	679	6	CD366811 full-1eng
12	627	33.3	691	5	BU682972 full-1eng
13	626	33.2	1051	5	BM545490 full-1eng
14	616	33.2	905	4	BM910528 full-1eng
15	615	32.7	660	6	CA306521 full-1eng
16	613	32.6	947	4	BM910416 full-1eng
17	607	32.2	1027	5	BM910416 full-1eng
18	604	32.1	655	7	BM910416 full-1eng
19	603	32.0	634	4	BM910416 full-1eng
20	600	31.9	1089	5	BM925673 full-1eng
21	600	31.8	1089	4	BM925673 full-1eng
22	599	31.8	1028	5	BM928759 full-1eng
23	599	31.8	1028	5	BM928759 full-1eng
24	598	31.8	732	6	CB321808 full-1eng

25	594	31.5	1131	2	BE907425
26	588	31.2	927	4	BG251151
27	586	31.1	800	4	BG685173
28	576	30.6	667	5	BM979999
29	576	30.6	1051	5	BM979999
30	569	30.2	571	7	CK819031
31	564	30.0	582	6	CA416835
32	564	30.0	644	7	BM973290
33	564	30.0	718	4	BM117988
34	559	29.7	638	5	BM688308
35	558	29.6	918	5	BM919631
36	555	29.5	588	2	AW245401
37	545	28.9	817	6	CB990003
38	535	28.4	780	6	CB994488
39	532	28.3	533	6	BE674778
40	531	28.2	831	6	CB992627
41	530	28.1	586	5	BP343733
42	522	27.7	582	5	BP253687
43	521	27.7	559	5	BP306442
44	517	27.5	832	4	BM827768
45	515	27.3	580	5	BP211545
46	515	27.3	793	2	BE883568
47	513	27.2	631	6	CD365564
48	509	27.0	931	5	BQ923358
49	508	27.0	654	2	BE391120
50	506	26.9	506	2	BF665646
51	502	26.7	678	7	BM973294
52	499	26.5	671	7	BM973287
53	490	26.0	931	5	BM973287
54	489	26.0	824	6	CB989607
55	485	25.8	576	4	BM126446
56	484	25.7	933	5	BM911774
57	483	25.7	783	5	BM602002
58	481	25.5	938	5	BM976554
59	476	25.3	826	6	CB991766
60	475	25.2	577	7	BM973291
61	472	25.1	574	7	CK819032
62	470	25.0	583	5	BP352691
63	469	24.9	582	5	BP360582
64	468	24.9	475	2	AW245758
65	466	24.7	878	4	BM976130
66	463	24.6	567	6	BM837565
67	461	24.5	527	4	BM547754
68	460	24.4	684	7	BM973289
69	458	24.3	1428	5	BM906070
70	454	24.1	1033	5	BM927715
71	452	24.0	1043	5	BM918017
72	449	23.8	799	4	BG681575
73	448	23.8	684	7	BM973289
74	446	23.6	837	4	BM976987
75	444	23.6	485	7	CR746989
76	443	23.5	680	6	CB530011
77	442	23.5	479	2	BE300872
78	442	23.5	480	2	BE300872
79	441	23.4	570	4	BG284887
80	440	23.4	1042	5	BQ072022
81	436	23.2	657	7	BM973295
82	435	23.1	680	4	BM976385
83	432	22.9	432	2	AW270125
84	429	22.8	583	5	BP305292
85	428	22.7	479	7	BM973293
86	427	22.7	581	5	BP333294
87	426	22.6	891	5	BM976385
88	424	22.5	432	1	BM976385
89	424	22.5	1313	5	BQ709748
90	420	22.3	842	4	BM976385
91	416	22.1	513	4	BM976385
92	415	22.1	1151	5	BM976385
93	412	21.9	486	5	BM976385
94	394	20.9	408	5	BM976385
95	393	20.9	402	7	BM976385
96	390	20.7	485	1	BM976385
97	389	20.7	447	4	BM857772

98 388 20.6 623 4 BG819012 602781053
 c 99 386 20.5 1166 5 BX342836 BX342836
 c 100 385 20.4 853 1 AL519409 AL519409

ALIGNMENTS

RESULT 1
 CR619301 1866 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0DC015J24 of Neuroblastoma Cot
 DEFINITION 25-normalized of Homo sapiens (human).
 ACCESSION CR619301
 VERSION CR619301.1 GI:50500108
 KEYWORDS HTC; cNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
 REMARK Full-length cDNA libraries and normalization

REFERENCE
 AUTHORS Unpublished
 TITLE Contact : Feng Liang Email : fliang@lifetech.com URL :
 JOURNAL http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 REMARK Faraday Avenue
 2 (bases 1 to 1866)
 Genoscope.
 REFERENCE Direct Submission
 AUTHORS Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)

COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 source Location/Qualifiers
 1. 1866
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC015J24"
 /issue_type="Neuroblastoma Cot 25-normalized"
 /plasmid="pCMVSPORT_6"

ORIGIN

Query Match 57.5%; Score 1082; DB 3; Length 1866;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 49 CGGCGGTGCGCGGAGTATCGAGCGCGAGATCGCGGGTTCGAGCTGTCCCG 108
 DB 1 CGGCGGTGCGCGGAGTATCGAGCGCGAGATCGCGGGTTCGAGCTGTCCCG 60
 QY 109 CTCCTCTCCCGGGAATGCGCTATCGAGTCCGCGCCCGCGTCTGAGAGAGAT 168
 DB 61 CTCCTCTCCCGGGAATGCGCTATCGAGTCCGCGCCCGCGTCTGAGAGAGAT 120
 QY 169 GAGTGGCTGAGTACTACGAGATGCTGCTTCAACCGTATGTTCAAGTGTGGCGGG 228
 DB 121 GAGTGGCTGAGTACTACGAGATGCTGCTTCAACCGTATGTTCAAGTGTGGCGGG 180
 QY 229 CAATGACCGAGTGCAGCTGAGCTCCGCGCTTCTGTGATAGAGCTCTCGCGCGC 288
 DB 181 CAATGACCGAGTGCAGCTGAGCTCCGCGCTTCTGTGATAGAGCTCTCGCGCGC 240
 QY 289 GCGGAGGCTTACCGCGCGCGCGCGAGCGGCTAGAGCTCTGCTGAGCTGAGCGCCGC 348
 DB 241 GCGGAGGCTTACCGCGCGCGCGCGAGCGGCTAGAGCTCTGCTGAGCTGAGCGCCGC 300
 QY 349 GGGCAGTGGCGGAGAGCAACTGCGGCTGCTGCGGCAACTCTCTGCGCGTGTGCGCCGC 408

DB 301 GGGCAGTGGCGGAGAGCAACTGCGGCTGCTGCGGCAACTCTCTGCGCGTGTGCGCCGC 360
 QY 409 CACGACTGCTGCGCGCACCTGCGCGGCAAGCGCGCCCGCGCAGTGTCTCAAGACGCTAT 468
 DB 361 CACGACTGCTGCGCGCACCTGCGCGGCAAGCGCGCCCGCGCAGTGTCTCAAGACGCTAT 420
 QY 469 AGCTATGGCACTTCCAGCTCTTCAAGAGACAGAGGCTAGTCCGCTGCGCTGCGCAG 528
 DB 421 AGCTATGGCACTTCCAGCTCTTCAAGAGAGACAGAGGCTAGTCCGCTGCGCTGCGCAG 480
 QY 529 TCAAGCAAGTTCGCAAAATTCAGAGAGGTCAGTGGAGACAGAGTCCCGCCCAACCA 588
 DB 481 TCAAGCAAGTTCGCAAAATTCAGAGAGGTCAGTGGAGACAGAGTCCCGCCCAACCA 540
 QY 589 CGGACGCGCGGAGTGGGCGCGCGCGCAAGTGTGTCAGACGCGCGCGAGAGAGGCGC 648
 DB 541 CGGACGCGCGGAGTGGGCGCGCGCGCAAGTGTGTCAGACGCGCGCGAGAGAGGCGC 600
 QY 649 CCAGCCGCAACCCAGCAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAGTAC 708
 DB 601 CCAGCCGCAACCCAGCAGCAGTCAAGAGCCCGCAGACCTTCTCTGAAGGCAAGTAC 660
 QY 709 TGTGACATCCGGCTCCGGTTGAGACAGATCTGCGAGATGGGCGAGCTTGGAGAC 768
 DB 661 TGTGACATCCGGCTCCGGTTGAGACAGATCTGCGAGATGGGCGAGCTTGGAGAC 720
 QY 769 GCGTGGCATCCCGCGCGCGCGCGCGCTGCGCGCGCAGCTGAGCGTGTGGGAGGCGC 828
 DB 721 GCGTGGCATCCCGCGCGCGCGCGCGCGCTGCGCGCGCAGCTGAGCGTGTGGGAGGCGC 780
 QY 829 ACCGAGTGTCTGCGCTCAAGGAGCTTGGGCTTGTGATCAATCAAGTTCTCAGAG 888
 DB 781 ACCGAGTGTCTGCGCTCAAGGAGCTTGGGCTTGTGATCAATCAAGTTCTCAGAG 840
 QY 889 CTCTCTATCTGAGAGGCTTCTGCGGCGCATCTAGTGGGCGCTGTGAGGCGCTG 948
 DB 841 CTCTCTATCTGAGAGGCTTCTGCGGCGCATCTAGTGGGCGCTGTGAGGCGCTG 900
 QY 949 CGGCGGTGCTCTGATGAGGCGCTGCGAGAGGCTGTGGCGCGGAGGCTTGGCGCTG 1008
 DB 901 CGGCGGTGCTCTGATGAGGCGCTGCGAGAGGCTGTGGCGCGGAGGCTTGGCGCTG 960
 QY 1009 CTGTGATGTGTGATGAGGCTGATGATGAGCTGTGCGCGCGCTGTGCTATGAG 1068
 DB 961 CTGTGATGTGTGATGAGGCTGATGATGAGCTGTGCGCGCGCTGTGCTATGAG 1020
 QY 1069 GAGGAAGGGGGGCGCGCGCGCGCAGAGGCTCTGATCCAGGATGGCAGATTATCC 1128
 DB 1021 GAGGAAGGGGGGCGCGCGCGCGCAGAGGCTCTGATCCAGGATGGCAGATTATCC 1080
 QY 1129 ACCTCCAGTCTCGGGCACCTTCTCTGAGAGACGACATCTTACCTT 1181
 DB 1081 ACCTCCAGTCTCGGGCACCTTCTCTGAGAGACGACATCTTACCTT 1133

RESULT 2
 CR593642 1860 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0DK010YK16 of HeLa cells Cot 25-normalized
 DEFINITION of Homo sapiens (human).
 ACCESSION CR593642
 VERSION CR593642.1 GI:50474449
 KEYWORDS HTC; cNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.
 REMARK Full-length cDNA libraries and normalization
 Unpublished
 Contact : Feng Liang Email : fliang@lifetech.com URL :

	Matches	1097	Conservative	0	Mismatches	1	Indels	0	Gaps	0
QY	84	TGCGCGGGATTCTGAGCTTTTCGCGCTTCCTCCCTCCCGGGAAATGGCGCTATTCGGGTGCAC	143							
Db	49	TGGCGGGATTCTGAGCTTTTCGCGCTTCCTCCCTCCCGGGAAATGGCGCTATTCGGGTGCAC	108							
QY	144	CCCGAGCCCGGCTCGGGAGAGAGATGAGTGTGCTGGAATACTAACAGAGATGCTGTGCCTTCA	203							
Db	109	CCCGAGCCCGGCTCGGGAGAGAGATGAGTGTGCTGGAATACTAACAGAGATGCTGTGCCTTCA	168							
QY	204	CCGATATGTTGAGGTGTGTGGCGGGCAACTGACCGAGTGCAGAGCTTGAGCTTCGGCCTT	263							
Db	169	CCGATATGTTGAGGTGTGTGGCGGGCAACTGACCGAGTGCAGAGCTTGAGCTTCGGCCTT	228							
QY	264	TCTGTGTGATTAAGGCTCTTGCGCGCCCGCGAGGCTTAAGCTCCGGGCGCGCAAGGCTTACA	323							
Db	229	TCTGTGTGATTAAGGCTCTTGCGCGCCCGCGAGGCTTAAGCTCCGGGCGCGCAAGGCTTACA	288							
QY	324	GCTCCTGTGTGAGCTGAGACGCGCCGGGCAATGTGCGGCGAGAGCAACTGCGGCTGTGGG	383							
Db	289	GCTCCTGTGTGAGCTGAGACGCGCCGGGCAATGTGCGGCGAGAGCAACTGCGGCTGTGGG	348							
QY	384	GCAACTCTCTGCGCGTGTGTGGCCCGGCAAGACTGTGCGGCAACTGTGCGCGCAAGCGCG	443							
Db	349	GCAACTCTCTGCGCGTGTGTGGCCCGGCAAGACTGTGCGGCAACTGTGCGCGCAAGCGCG	408							
QY	444	CCGGCCAGTGTCTCCAGAACGCTAATGCTAATGGCACTCTCCAGCTTCTTAAGAGAGACAGA	503							
Db	409	CCGGCCAGTGTCTCCAGAACGCTAATGCTAATGGCACTCTCCAGCTTCTTAAGAGAGACAGA	468							
QY	504	GGGTATGCTGCGGTGCGCGGTGCGGCAAGTATTCGCAATTCATGACAGGGTCAAGTG	563							
Db	469	GGGTATGCTGCGGTGCGCGGTGCGGCAAGTATTCGCAATTCATGACAGGGTCAAGTG	528							
QY	564	GGAAGCAAGGCTCCCCCCCAACCAAGCGGCAAGCGGAGTGTGGGGCGGCCCAAGTGTGG	623							
Db	529	GGAAGCAAGGCTCCCCCCCAACCAAGCGGCAAGCGGAGTGTGGGGCGGCCCAAGTGTGG	588							
QY	624	TGCCAGACGCGCGGCGAGAGGGGCCCCAGCGGCAACCCAGCAACATGATCAAGCCCGCCAG	683							
Db	589	TGCCAGACGCGCGGCGAGAGGGGCCCCAGCGGCAACCCAGCAACATGATCAAGCCCGCCAG	648							
QY	684	ACCTTCCCTCTAAGGCAAAATGATCCTGTGACATCCGCGTCCGGGTTGAGAGAGATGACTG	743							
Db	649	ACCTTCCCTCTAAGGCAAAATGATCCTGTGACATCCGCGTCCGGGTTGAGAGAGATGACTG	708							
QY	744	CGAGCATGGGCGACGCTTGAGAGACAGGGGCTGGGCAATCCGCGGCGCCCAAGCGCTGACGCG	803							
Db	709	CGAGCATGGGCGACGCTTGAGAGACAGGGGCTGGGCAATCCGCGGCGCCCAAGCGCGTGGCGCG	768							
QY	804	GCACTGAGACGTGTTTGGGCGAGGCCCAACCGCAGGTGTGCGTCTAAGAGAACTTGGGCTGTGT	863							
Db	769	GCACTGAGACGTGTTTGGGCGAGGCCCAACCGCAGGTGTGCGTCTAAGAGAACTTGGGCTGTGT	828							
QY	864	GCTTTGTGACATCAAGTCTCAGAGCTCTCCTAATGTGAAACCTTTGTGGGCGACTAAGCT	923							
Db	829	GCTTTGTGACATCAAGTCTCAGAGCTCTCCTAATGTGAAACCTTTGTGGGCGACTAAGCT	888							
QY	924	GAGTGGCGCCCTGCTGACAGGCGCCCTGCGGGGCGGTGTTCTGACTGAGGCGCTTGCAGAGGC	983							
Db	889	GAGTGGCGCCCTGCTGACAGGCGCCCTGCGGGGCGGTGTTCTGACTGAGGCGCTTGCAGAGGC	948							
QY	984	TGTGGGCGGGAGAGGCTGTTCCGCTTGTGATGTGATGAGGCTGATGAGAGCTTG	1043							
Db	949	TGTGGGCGGGAGAGGCTGTTCCGCTTGTGATGTGATGAGGCTGATGAGAGCTTG	1008							
QY	1044	CCGGCGCGGCTGTGTGCTGATGAGAGAGAGAGGGGGCGGGCCCGACAGAGGCTCTTG	1103							
Db	1009	CCGGCGCGGCTGTGTGCTGATGAGAGAGAGAGGGGGCGGGCCCGACAGAGGCTCTCTG	1068							
QY	1104	ATCCAGAGACTGGCAGAGATTGAATCCCAACTTCAAAGTCTCCGGGCAACTTCTCTGGGAGG	1163							
Db	1069	ATCCAGAGACTGGCAGAGATTGAATCCCAACTTCAAAGTCTCCGGGCAACTTCTCTGGGAGG	1128							

Y	1164	ACGACCATCTTACCCCT	1181
Db	1129	ACGACCATCTTACCCCT	1146
RESULT 4			
LOCUS	CR604307		
DEFINITION	CR604307	1053 bp	mRNA
ACCESSION	CR604307		
VERSION	CR604307.1		
KEYWORDS	HTC; cDNA		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Li, W. B., Gruber, C., Jesse, J., and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished		
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue		
REFERENCE	Genoscope.		
AUTHORS	Genoscope.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		
FEATURES	Location/Qualifiers		
source	1..1053		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CS0DU006TL09"		
	/tissue_type="B cells (Ramos cell line) Cot 25-normalized"		
	/plasmid="pCMVSPORT_6"		
ORIGIN			
Query Match	53.2%; Score 1002; DB 3; Length 1053;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1052; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Y	26	GAAAGATCCAAACCAAGTGGCTGGCGCGCTGCGCCAGAGATCATCGAGCGCAGAAATCTG	85
Db	1	GAAAGATCCAAACCAAGTGGCTGGCGCGCTGCGCCAGAGATCATCGAGCGCAGAAATCTG	60
Y	86	GCCGGGTTCTGAGCTTGTTCGCTCCCTCCCGGGGAATGCGCTATCCGGGTTCGACCC	145
Db	61	GCCGGGTTCTGAGCTTGTTCGCTCCCTCCCGGGGAATGCGCTATCCGGGTTCGACCC	120
Y	146	CGGCGCCCGGTGGTGGAGAGAGATGAGTGGCTCGAGCTACTACGGGATGCTGTCCGCTTACAC	205
Db	121	CGGCGCCCGGTGGTGGAGAGAGATGAGTGGCTCGAGCTACTACGGGATGCTGTCCGCTTACAC	180
Y	206	GTATGTTTGAAGTGTGGTGGCGGCGCACTGACCGAGTCCGAGCTGGAAGCTCTGTGAGCTTTTC	265
Db	181	GTATGTTTGAAGTGTGGTGGCGGCGCACTGACCGAGTCCGAGCTGGAAGCTCTGTGAGCTTTTC	240
Y	266	TGCTGATGAGGCTCTCTGGCGCCCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGAC	325
Db	241	TGCTGATGAGGCTCTCTGGCGCCCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGAC	300
Y	326	TTCCTGCTGAGGCTGAGAGCGCGCGGAGGAGTGGCGGCGAGAGCAACTTGCGGCTGTGCTGAGGC	385
Db	301	TTCCTGCTGAGGCTGAGAGCGCGCGGAGGAGTGGCGGCGAGAGCAACTTGCGGCTGTGCTGAGGC	360

```

QY 386 AACCTCTGCGCGTCTGCGCCGCGCAGCACTGTGCGCAGCTGCGCGCGCGCGCC 445
Db 361 AACTCTGCGCGTCTGCGCCGCGCAGCACTGTGCGCGCAGCTGCGCGCGCGCGCC 420
QY 446 GGCAGTGTCTCCAGAACGCTATAGCTATGTGACCTTCAGCTCTTCAAGAGAGACAGAG 505
Db 421 GGCAGTGTCTCCAGAACGCTATAGCTATGTGACCTTCAGCTCTTCAAGAGAGACAGAG 480
QY 506 GTAGCTGCGCTGCGCGCGCGCAGTCAAGAGAGCTTCTGCAATTTCTGACAGAGTCAAGTGG 565
Db 481 GTAGCTGCGCTGCGCGCGCGCAGTCAAGAGAGCTTCTGCAATTTCTGACAGAGTCAAGTGG 540
QY 566 AGACAGAGCTTCTGCGCGCGCGCAGTCAAGAGAGCTTCTGCAATTTCTGACAGAGTCAAGTGG 625
Db 541 AGACAGAGCTTCTGCGCGCGCGCAGTCAAGAGAGCTTCTGCAATTTCTGACAGAGTCAAGTGG 600
QY 626 CCAGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
Db 601 CCAGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 686 CTTCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 745
Db 661 CTTCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 746 AGCATGCGCGCGCGCTTGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 805
Db 721 AGCATGCGCGCGCGCTTGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 806 AGCTGAGAGCTGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 865
Db 781 AGCTGAGAGCTGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 866 TTGTGATCATCATGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 925
Db 841 TTGTGATCATCATGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 926 GTGCGCGCGCGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 985
Db 901 GTGCGCGCGCGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 986 TGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1045
Db 961 TGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1046 GGGCGCGCGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
Db 1021 GGGCGCGCGCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1053

```

```

RESULT 5
LOCUS BM908148 1062 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6707554 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5745179
5', mRNA sequence.
ACCESSION BM908148
VERSION BM908148.1 GI:19358527
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1062)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: c9abrs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

```

```

FEATURES
Source
1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5745179"
/issue_type="medulla"
/lab_host="DH10B"
/clone_id="NIH_MGC_119"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invited). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 38.8%; Score 731; DB 5; Length 1062;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 731; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 GTGTCTCCAGAACGCTATAGCTATGCGACCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
Db 73 GTGTCTCCAGAACGCTATAGCTATGCGACCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132
QY 511 TGGCGTGGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
Db 133 TGGCGTGGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 192
QY 571 GGTCTCCCGCCCAACCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630
Db 193 GGTCTCCCGCCCAACCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 252
QY 631 CGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 690
Db 253 CGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 691 TCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
Db 313 TCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
QY 751 GGGCGAGCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 810
Db 373 GGGCGAGCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
QY 811 GAGGTGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
Db 433 GAGGTGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 871 GACATCAATTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 930
Db 493 GACATCAATTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
QY 931 GGCCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
Db 553 GGCCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612
QY 991 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1050
Db 613 CGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
QY 1051 CGCTGTTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1110
Db 673 CGCTGTTTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
QY 1111 ACTGAGAGATGATCCAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1170

```

```
Db 733 ACTGGAGGATGATCCACCTCCAAAGTCCGGGCGCACCTTCTCTGGAGGACGACCA 792
Qy 1171 TCTCTACCCCT 1181
Db 793 TCTCTACCCCT 803

RESULT 6
CR59685
LOCUS full-length cDNA clone CS0D1069YH17 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human) .
ACCESSION CR59685
VERSION CR59685.1 GI:50480492
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1510)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1510)
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 1510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1069YH17"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

FEATURES
source

ORIGIN
Query Match 38.8%; Score 731; DB 3; Length 1510;
Best Local Similarity 100.0%; Pred.No.0;
Matches 731; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 451 GTGTCTCCAGACGCTATAGCTATGGACCTCCAGCTCTTCAAGAGGAGACAGGGTAGC 510
Db 29 GTGTCTCCAGACGCTATAGCTATGGACCTCCAGCTCTTCAAGAGGAGACAGGGTAGC 88
Qy 511 TGGCGTCCGCGTCCGAGTCAAGCAAGTCTGCAAAATCTTCAGAGAGGTCAGTGGAGACA 570
Db 89 TGGCGTCCGCGTCCGAGTCAAGCAAGTCTGCAAAATCTTCAGAGAGGTCAGTGGAGACA 148
Qy 571 GGGTCCCCCAACCAAGAGGCGGCGGAGTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 630
Db 149 GGGTCCCCCAACCAAGAGGCGGCGGAGTCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 208
Qy 631 CGGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 690
Db 209 CGGCGGCGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 268
Qy 691 TCTGAAGGCAAGTGAAGTCTGATCATCCGCTCCGGGTTTCAGACAGAGTACTGCGAGCAT 750
Db 269 TCTGAAGGCAAGTGAAGTCTGATCATCCGCTCCGGGTTTCAGACAGAGTACTGCGAGCAT 328
Qy 751 GGGCGACGCTTTGAGACAGGGCGGTGGCATCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 810
```

```
Db 329 GGGCGACGCTTGGAGCAGGAGGCGGTGCATCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 388
Qy 811 GACGTGTTTGGCAGAGCCACCGCAGTGCCTGCGCTCAAGGACCTTGCGTTTGT 870
Db 389 GACGTGTTTGGCAGAGCCACCGCAGTGCCTGCGCTCAAGGACCTTGCGTTTGT 448
Qy 871 GACATCAAGTTTTCAGAGCTCTCTATCTGACGCGCTTCTGGGGCGCATCTAGTGGC 930
Db 449 GACATCAAGTTTTCAGAGCTCTCTATCTGACGCGCTTCTGGGGCGCATCTAGTGGC 508
Qy 931 GCGCTGCTGACAGCGCTGCGGGCGGTGTTCTGACTGAGAGCCCTGCGAGAGGCTGAGGC 990
Db 509 GCGCTGCTGACAGCGCTGCGGGCGGTGTTCTGACTGAGAGCCCTGCGAGAGGCTGAGGC 568
Qy 991 CGGAGGCTGTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
Db 569 CGGAGGCTGTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628
Qy 1051 CGCTGTTGCTGATGAGAGGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1110
Db 629 CGCTGTTGCTGATGAGAGGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 688
Qy 1111 ACTGGAGGATGATCCCAAGTCTCCGGGCGACCTTCTGCTGGAGAGACCA 1170
Db 689 ACTGGAGGATGATCCCAAGTCTCCGGGCGACCTTCTGCTGGAGAGACCA 748
Qy 1171 TCTCTACCCCT 1181
Db 749 TCTCTACCCCT 759

RESULT 7
CR590827
LOCUS full-length cDNA clone CS0DE001YM11 of Placenta of Homo sapiens
DEFINITION (human) .
ACCESSION CR590827
VERSION CR590827.1 GI:50471634
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1586)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1586)
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1. 1586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE001YM11"
/tissue_type="Placenta"
/plasmid="pCMVSPORT_6"

FEATURES
source

ORIGIN
Query Match 38.8%; Score 731; DB 3; Length 1586;
```



```

QY      623 GTGCCAGACGGCGCGAGAGGGGCCCCAGCCGCGACCCGACAGCAAGTCAGAGCCCCGCCA 662
Db      601 GTCCAGACGGCGCGAGAGGGGCCCCAGCCGCGACCCGACAGCAAGTCAGAGCCCCGCCA 660
QY      683 GACCTTCCTTGAAAGGAAGTGAACCTGTGACATCCGGCTCCGGGTTCCAGAGAGATGACT 742
Db      661 GACCTTCCTTGAAAGGAAGTGAACCTGTGACATCCGGCTCCGGGTTCCAGAGAGATGACT 720
QY      743 GCGAGCATGGCGCAGCCTTGAGACAGAGCGCTGGCATGCC 781
Db      721 GCGAGCATGGCGCAGCCTTGAGACAGAGCGCTGGCATGCC 759

RESULT 9
BU902131      1161 bp      mRNA      linear      EST 17-OCT-2002
LOCUS      AGENCOURT_10127738 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502805
DEFINITION      5', mRNA sequence.
ACCESSION      BU902131
VERSION      BU902131.1 GI:24084044
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabs-remail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Life Technologies, Inc.
      DNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LLAM14059 row: h column: 06
      High quality sequence stop: 707.
      Location/Qualifiers
        1..1161
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:6502805"
        /tissue_type="fetal yolk sac"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NciI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
        Average insert size 2.1 kb."
ORIGIN
Query Match      36.1%; Score 680; DB 5; Length 1161;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 730; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      451 GTGCTTCACAAAGCTTACCTATGAGCACTCCAGCTCTTCAAGAGAGAGAGGGTAC 510
Db      77 GTGCTTCACAAAGCTTACCTATGAGCACTCCAGCTCTTCAAGAGAGAGAGGGTAC 136
QY      511 TGGCGTGGCGGTGGCGAGTCAAGAGTTCTGCAAAATCTCAGCAGGGTCACTGGGAGACA 570
Db      137 TGGCGTGGCGGTGGCGAGTCAAGAGTTCTGCAAAATCTCAGCAGGGTCACTGGGAGACA 136
QY      571 GGGTCCCCCAACCAAGCGGAGCGGAGAGTCCGGGCGCGCCCACTGTGTGCCAGA 630
Db      197 GGGTCCCCCAACCAAGCGGAGCGGAGAGTCCGGGCGCGCCCACTGTGTGCCAGA 256
QY      631 CGGCGGCGGAGAGGGGCCCCAGCCGACCCGACAGAGATCAAGCCCGCAACCTTCC 690

```

```

Db      257 CGGCGGCGGAGAGGGGCCCCAGCCGCGACCCGACAGCAAGTCAGAGCCCCGCCA 316
QY      691 TCTGAAGGCAAAAGTGAACCTGTGACATCCGGCTCCGGGTTGAGAGAGATGTCGAGCAT 750
Db      317 TCTGAAGGCAAAAGTGAACCTGTGACATCCGGCTCCGGGTTGAGAGAGATGTCGAGCAT 376
QY      751 GGGCAGCCTTGGAGAGAGAGGCGTGGATCCCGGCGGCGCCAGAGCGTGGCGGAGAGCTG 810
Db      377 GGGCAGCCTTGGAGAGAGGCGTGGATCCCGGCGGCGCCAGAGCGTGGCGGAGAGCTG 436
QY      811 GACGTGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 870
Db      437 GACGTGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
QY      871 GACATCAAGTTCTCAGAGCTCTTATCTGACAGCCCTTGGAGAGAGAGAGAGAGAGAGAG 930
Db      497 GACATCAAGTTCTCAGAGCTCTTATCTGACAGCCCTTGGAGAGAGAGAGAGAGAGAGAG 556
QY      931 GCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
Db      557 GCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616
QY      991 CGGAGAGCTGTTGCGCTGCTGAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1050
Db      617 CGGAGAGCTGTTGCGCTGCTGAGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 676
QY      1051 CGCCTGTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1110
Db      677 CGCCTGTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736
QY      1111 ACTGAGAGATGATCCCACTTCCAAAGTCTCCGGGCGACCTTCTCTGGAGAGAGAGACA 1170
Db      737 ACTGAGAGATGATCCCACTTCCAAAGTCTCCGGGCGACCTTCTCTGGAGAGAGAGACA 796
QY      1171 TCTCTACCCCT 1181
Db      797 TCTCTACCCCT 807

RESULT 10
BM559141      993 bp      mRNA      linear      EST 20-FEB-2002
LOCUS      AGENCOURT_5562852 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5551258
DEFINITION      5', mRNA sequence.
ACCESSION      BM559141
VERSION      BM559141.1 GI:18802480
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE      Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabs-remail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Life Technologies, Inc.
      DNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: LLAM12264 row: 1 column: 11
      High quality sequence start: 31
      High quality sequence stop: 611.
      Location/Qualifiers
        1..993
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5551258"
FEATURES
source

```

/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="N1H MG6 67"
/note="Organ: eye; Vector: PCMV-SPORE6; Site:1: NCI;
Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 33.4%; Score 629; DB 4; Length 993;
Best Local Similarity 99.9%; Pred. No. 3.2e-309; Indels 0; Gaps 0;
Matches 679; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 502 GAGGGTAGTCCGCGCCGCGCGAGTCAAGCAATTCTGCAATTCAGCAGGCTCAG 561
DB 41 GAGGGTAGTCCGCGCCGCGCGAGTCAAGCAATTCTGCAATTCAGCAGGCTCAG 100
QY 562 TGGGAGCAGGCTCCCGCCCAACCAAGCGGAGCGGAGTCCGGGCGCGCCAGTGT 621
DB 101 TGGGAGCAGGCTCCCGCCCAACCAAGCGGAGCGGAGTCCGGGCGCGCCAGTGT 160
QY 622 GGTGCCAGCGCGCGGAGAGGGGCGCGAGCGCGCCAGCCCAAGCAGTCAAGCCGCG 681
DB 161 GGTGCCAGCGCGCGGAGAGGGGCGCGAGCGCGCCAGCCCAAGCAGTCAAGCCGCG 220
QY 682 AGACCTTCTCTGAAGCAAGTACCTGTGACATCCGCGCTCCGGGTTCCAGCAGAGTAC 741
DB 221 AGACCTTCTCTGAAGCAAGTACCTGTGACATCCGCGCTCCGGGTTCCAGCAGAGTAC 280
QY 742 TCGGAGCATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 801
DB 281 TCGGAGCATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 340
QY 802 CGGAGCTGAGCGGTTTGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 861
DB 341 CGGAGCTGAGCGGTTTGGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 400
QY 862 GTGGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGAGCGGCTTCTGGGCGGCTAC 921
DB 401 GTGGTTGTGACATCAAGTCTCAGAGCTCTCTATCTGAGCGGCTTCTGGGCGGCTAC 460
QY 922 CTGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 981
DB 461 CTGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 520
QY 982 GCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1041
DB 521 GCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 580
QY 1042 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1101
DB 581 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 640
QY 1102 TGATCCAGAGCTGGCAGAGATTGATCCCACTCCAGTCTCCGGGCGGCGGCGG 1161
DB 641 TGATCCAGAGCTGGCAGAGATTGATCCCACTCCAGTCTCCGGGCGGCGGCGGCG 700
QY 1162 GGAAGACCATCTCAACCCCT 1181
DB 701 GGAAGACCATCTCAACCCCT 720

RESULT 11

CD366811/c 679 bp mRNA linear EST 05-AUG-2004
LOCUS
DEFINITION
UI-H-FT2-bjo-p-06-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
UI-H-FT2-bjo-p-06-0-UI 3', mRNA sequence.

ACCESSION
CD366811
VERSION
CD366811.1 GI:31150901
KEYWORDS
EST.
Homo sapiens (human)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 679)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Straubeberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.iowa.edu/distribution/cgap.html>
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

1. 679
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_id="N1H CGAP FT2"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI-CGAP FT2 is a subcloned cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The mRNA samples
were a mixture of these conditions (times refer to
incubations following isolation by bronchoalveolar lavage)
(some normal donor macrophages were cultured in some of
the conditions, other donor macrophages in different
conditions). The mRNA samples were pooled for library
construction. Control 0 hours; control 3 hours; control 24
hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours;
PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella
moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph
aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours;
Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours;
Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt
adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24
hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours;
wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours.
The library was subcloned according to Bonaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. The tissue
was provided by Dr. Gary W. Hunninghake of the University
of Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LTB=UI-H-FT2
TAG_SEQ=GAGCATGCGG"

ORIGIN

Query Match 33.3%; Score 627; DB 6; Length 679;
Best Local Similarity 100.0%; Pred. No. 3.3e-308; Indels 0; Gaps 0;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1257 TACCCCAAGTGTGCCCCCTTAGCCTCCGAGGCGGGGCGTGGGCTGTATCTCAGAAAG 1316
DB 645 TACCCCAAGTGTGCCCCCTTAGCCTCCGAGGCGGGGCGTGGGCTGTATCTCAGAAAG 586
QY 1317 GAGGGGCAAGCTACACACTCAACCAAGGCCCCCTGACATTTATCTGTATCTTTGGG 1376
DB 585 GAGGGGCAAGCTACACACTCAACCAAGGCCCCCTGACATTTATCTGTATCTTTGGG 526
QY 1377 CTGTGTGCACTGTCAAGGTCACACACTCGCTCATGTCTCAACATGCGCCCTGCTGAGATC 1436
DB 525 CTGTGTGCACTGTCAAGGTCACACACTCGCTCATGTCTCAACATGCGCCCTGCTGAGATC 466
QY 1437 TTCCCTGGGCGCTTGCGCTGCGCTGCTTCCAGACACACTTCTTGGCTTAAAGGCTTC 1496

Db 465 TTCCTGGGCTCTGCTGCTGGCTGCTGCCAGACACACTTCTTGGCTTAAGGGCTTC 406
QY 1497 TCTCTCAGAGACCTCTTAATTGACACCAACCAACTGGGCTTACGCCATCATGAGGGCAG 1556
Db 405 TCTCTCAGAGACCTCTTAATTGACACCAACCAACTGGGCTTACGCCATCATGAGGGCAG 346
QY 1557 TGGAGCTGGGGGTGCACATGGGGGCTGCTCACTTGGCCACACATCTTCACAGCCAGGG 1616
Db 345 TGGAGCTGGGGGTGCACATGGGGGCTGCTCACTTGGCCACACATCTTCACAGCCAGGG 286
QY 1617 CCTGCGCCAGCTTCAATTTAAGACCTGACTCCCTCCACCTTCCCGCTGCTCCAGAG 1676
Db 285 CCTGCGCCAGCTTCAATTTAAGACCTGACTCCCTCCACCTTCCCGCTGCTCCAGAG 226
QY 1677 CTGAACATGACTTGTGACTTGTGATGTCACTGGAGTGTCACTAGGAGTGTATGAGCAGC 1736
Db 225 CTGAACATGACTTGTGACTTGTGATGTCACTGGAGTGTCACTAGGAGTGTATGAGCAGC 166
QY 1737 ATCATACCAAGGCTTACTGTTGACATGGGGGCAAAACAGTAAACAGCCACTTCTTGG 1796
Db 165 ATCATACCAAGGCTTACTGTTGACATGGGGGCAAAACAGTAAACAGCCACTTCTTGG 106
QY 1797 AAGGGAATGCAGAGCTTGGGGGTGATGGAAGAACCTTTTACAATGATACCAATTA 1856
Db 105 AAGGGAATGCAGAGCTTGGGGGTGATGGAAGAACCTTTTACAATGATACCAATTA 46
QY 1857 AACTGCCCTGGAAGGCAATAGGTGGG 1883
Db 45 AACTGCCCTGGAAGGCAATAGGTGGG 19

RESULT 12
BU682972/c 691 bp mRNA linear EST 07-OCT-2002
LOCUS UI-CF-EC1-acl-g-05-0-UI s1 UI-CF-EC1 Homo sapiens cDNA clone
DEFINITION UI-CF-EC1-acl-g-05-0-UI 3', mRNA sequence.
ACCESSION BU682972 GI:23534401
VERSION BU682972.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 691)
AUTHORS Bonaldi,M.F., Lemmon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul.mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.research.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLI=yes
Location/Qualifiers
1. .691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acl-g-05-0-UI"
/tissue_type="Lung"

FEATURES
source

1. .691
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EC1-acl-g-05-0-UI"
/tissue_type="Lung"

/dev_stage="Adult and Fetal"
/lab_host="DHI08 (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EC1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polynucleotide linker; Site 1: EcoR I; Site 2: Not I;
UI-CF-EC1 is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldi, Lemmon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AAGCTTAC
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_LIB=UI-CF-EC1
TAG_SEQ=AAGTCTTAC"

ORIGIN

Query Match 33.3%; Score 627; DB 5; Length 691;
Best Local Similarity 100.0%; Pred.No. 3.3e-308;
Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1257 TACCCAGAGTGTGCCCCCTTACGCTCCGAGAGGGGGCTGGGCTGTATCTCAGAAAG 1316
Db 646 TACCCAGAGTGTGCCCCCTTACGCTCCGAGAGGGGGCTGGGCTGTATCTCAGAAAG 587
QY 1317 GAGGGGACAGCTTACACTACCAACCAAGGCCCTCTGACATTTGATCTCTGATCTTGGG 1376
Db 586 GAGGGGACAGCTTACACTACCAACCAAGGCCCTCTGACATTTGATCTCTGATCTTGGG 527
QY 1377 CTGTCTGACTGTGCACAGTGTGCACACACTGCTGATGCTCAGACTGCCCCGTGTGATC 1436
Db 526 CTGTCTGACTGTGCACAGTGTGCACACACTGCTGATGCTCAGACTGCCCCGTGTGATC 467
QY 1437 TTCCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1496
Db 466 TTCCCTGGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 407
QY 1497 TCTCTCAGAGACCTCTTAATTGACACCAACCAACTGGGCTTACGCCATCATGAGGGCAG 1556
Db 406 TCTCTCAGAGACCTCTTAATTGACACCAACCAACTGGGCTTACGCCATCATGAGGGCAG 347
QY 1557 TGGAGCTGGGGGTGCACATGGGGGCTGCTCACTTGGCCACACATCTTCACAGCCAGGG 1616
Db 346 TGGAGCTGGGGGTGCACATGGGGGCTGCTCACTTGGCCACACATCTTCACAGCCAGGG 287
QY 1617 CCTGCGCCAGCTTCAATTTAAGACCTGACTCCCTCCACCTTCCCGCTGCTCCAGAG 1676
Db 286 CCTGCGCCAGCTTCAATTTAAGACCTGACTCCCTCCACCTTCCCGCTGCTCCAGAG 227
QY 1677 CTGAACATGACTTGTGACTTGTGATGTCACTGGAGTGTCACTAGGAGTGTATGAGCAGC 1736
Db 226 CTGAACATGACTTGTGACTTGTGATGTCACTGGAGTGTCACTAGGAGTGTATGAGCAGC 167
QY 1737 ATCATACCAAGGCTTACTGTTGACATGGGGGCAAAACAGTAAACAGCCACTTCTTGG 1796
Db 166 ATCATACCAAGGCTTACTGTTGACATGGGGGCAAAACAGTAAACAGCCACTTCTTGG 107
QY 1797 AAGGGAATGCAGAGCTTGGGGGTGATGGAAGAACCTTTTACAATGATACCAATTA 1856
Db 106 AAGGGAATGCAGAGCTTGGGGGTGATGGAAGAACCTTTTACAATGATACCAATTA 47
QY 1857 AACTGCCCTGGAAGGCAATAGGTGGG 1883
Db 46 AACTGCCCTGGAAGGCAATAGGTGGG 20


```

RESULT 13
BX354590/c 1051 bp mRNA linear EST 08-APR-2004
LOCUS BX354590 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC022YK03 3-PRIME, mRNA sequence.
ACCESSION BX354590
VERSION BX354590.2 GI:46307295
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1051)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30377899.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9074.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DC022A02NP1&c=9074.r.
FEATURES
source
1. 1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC022YK03"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 33.2%; Score 626; DB 5; Length 1051;
Best Local Similarity 100.0%; Pred. No. 1.1e-307;
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1205 GGGCTGTGAGGCTTAACATTTTCAGCTGAGTTTCTTCCAGACTCTCTACCCCA 1264
DB 650 GGGCTGTGAGGCTTAACATTTTCAGCTGAGTTTCTTCCAGACTCTCTACCCCA 591
QY 1265 GGTGTGCCCCCTTAAGCTCCGGAAGGGGGGCTGTGCTTATCTCAAGAGGAGGCA 1324
DB 590 GGTGTGCCCCCTTAAGCTCCGGAAGGGGGGCTGTGCTTATCTCAAGAGGAGGCA 531
QY 1325 CAGCTACACACTACCAAGAGGCCCCCTGACATTTGATCTTGATTTGGGCTGTGCTG 1384
DB 530 CAGCTACACACTACCAAGAGGCCCCCTGACATTTGATCTTGATTTGGGCTGTGCTG 471
QY 1385 ACTGTACAGGTGACACACTGTGCTCATGTCACTGCCCCCTGTGAGATCTTCCCTGG 1444
DB 470 ACTGTACAGGTGACACACTGTGCTCATGTCACTGCCCCCTGTGAGATCTTCCCTGG 411
QY 1445 GCGTTCGCCCTGCGCTCTCCAGACACACTTTTGGCCTTAAGGGCTTCTCTCAG 1504
DB 410 GCGTTCGCCCTGCGCTCTCCAGACACACTTTTGGCCTTAAGGGCTTCTCTCAG 351
QY 1505 GACCTTAATTTTACCAACAACCTGTGCTTCAAGGACATAGTGGGCACTGAGCTG 1564
DB 350 GACCTTAATTTTACCAACAACCTGTGCTTCAAGGACATAGTGGGCACTGAGCTG 291
QY 1565 GGGTGCATGTGGGGCTGTCTACCTTGGCCACATCTCAGCCAGGAGGGCTTGGCC 1624
DB 1624 GGGTGCATGTGGGGCTGTCTACCTTGGCCACATCTCAGCCAGGAGGGCTTGGCC

```

```

DB 290 GGGTGCATGTGGGGCTGTCTACCTTGGCCACATCTCAGCCAGGAGGGCTTGGCC 221
QY 1625 AGCTTAATTTTACCAACAACCTGTGCTTCAAGGACATAGTGGGCACTGAGCTG 1684
DB 230 AGCTTAATTTTACCAACAACCTGTGCTTCAAGGACATAGTGGGCACTGAGCTG 171
QY 1685 AGACTTGCATTTGATGATGACCTGAGATGTCATCATGGAGGTGTTATGGCAGCATAC 1744
DB 170 AGACTTGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 111
QY 1745 AAGGCTTACTGTGTCACATGAGGCGCAAAACAGTAAACAGCCACTTCTTGAAGAGAA 1804
DB 110 AAGGCTTACTGTGTCACATGAGGCGCAAAACAGTAAACAGCCACTTCTTGAAGAGAA 51
QY 1805 TGCAGAGGCTTTGGGGGTGATGAGAA 1830
DB 50 TGCAGAGGCTTTGGGGGTGATGAGAA 25
RESULT 14
BI910528 905 bp mRNA linear EST 16-OCT-2001
LOCUS BI910528
DEFINITION 603068215F1 NIH_MGC_118 Homo sapiens CDNA clone IMAGE:5217218 5',
mRNA sequence.
ACCESSION BI910528
VERSION BI910528.1 GI:16173927
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 905)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1545 row: n column: 03
High quality sequence scop: 863.
FEATURES
source
1. 905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5217218"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC library."
ORIGIN
Query Match 32.7%; Score 616; DB 4; Length 905;
Best Local Similarity 100.0%; Pred. No. 1.4e-302;
Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 451 GTGTCTCCAGAAAGCTATAGTGGACCTTCAGCTTTCAAGAGGAGGAGGCTAC 510
DB 38 GTGTCTCCAGAAAGCTATAGTGGACCTTCAGCTTTCAAGAGGAGGAGGCTAC 97

```


QY 1689 TTGCACTTGATGTCACCTGAGTGTCACTGGAGTGTATGAGCATATACCAAG 1748
 DB 213 TTGCACTTGATGTCACCTGAGTGTCACTGGAGTGTATGAGCATATACCAAG 154
 QY 1749 CTTACTGTTGACATGAGGAGCCAAACAGTAAACAGCCACTTCTTGGAAGGAAATGCA 1808
 DB 153 CTTACTGTTGACATGAGGAGCCAAACAGTAAACAGCCACTTCTTGGAAGGAAATGCA 94
 QY 1809 AAGGCTTTGGGGGATGAGAAAGACCTTTTAAATAATGATCAATTAATCTGCGTGA 1868
 DB 93 AAGGCTTTGGGGGATGAGAAAGACCTTTTAAATAATGATCAATTAATCTGCGTGA 34
 QY 1869 AAGGCAATAGTGGG 1883
 DB 33 AAGGCAATAGTGGG 19
 RESULT 16
 BI910416 947 bp mRNA linear EST 16-OCT-2001
 LOCUS 603068087F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217213 5',
 DEFINITION mRNA sequence.
 ACCESSION BI910416 GI:16173804
 VERSION BI910416.1 GI:16173804
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 947)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L16M1545 row: m column: 22
 High quality sequence stop: 853.
 Location/Qualifiers
 1..947
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5217213"
 /issue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_118"
 /note="Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC library."
 ORIGIN
 Query Match 32.6%; Score 613; DB 4; Length 947;
 Best Local Similarity 100.0%; Pred. No. 4.8e-301;
 Matches 613; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 451 GTGTCTCCAGAAAGGCTATAGTATGACCTCCAGCTCTTCAAGAGAGAGAGGGTAC 510
 DB 72 GTGTCTCCAGAAAGGCTATAGTATGACCTCCAGCTCTTCAAGAGAGAGAGGGTAC 131

QY 511 TGGCGTGGCGCTGCGAGTCAAGACAGTCTGCAAAATTTCAAGAGAGGTCACTGGAGACA 570
 DB 132 TGGCGTGGCGCTGCGAGTCAAGACAGTCTGCAAAATTTCAAGAGAGGTCACTGGAGACA 191
 QY 571 GGTCTCCCTCCCAACCAAGGCGAGCGCGAGTCCGGGCCGCTGAGTGTGTGCAAGA 630
 DB 192 GGTCTCCCTCCCAACCAAGGCGAGCGCGAGTCCGGGCCGCTGAGTGTGTGCAAGA 251
 QY 631 CGGCGGAGAGAGGAGGAGCCCGAGCCGAGCCAGACAGTCAAGAGCCGAGACCTTCC 690
 DB 252 CGGCGGAGAGAGGAGGAGCCCGAGCCGAGCCAGACAGTCAAGAGCCGAGACCTTCC 311
 QY 691 TCTGAAGGCAAAAGTCACTGTGACATCCGAGTCCGGGTTTCAAGCAGAGTACTGCGAGAT 750
 DB 312 TCTGAAGGCAAAAGTCACTGTGACATCCGAGTCCGGGTTTCAAGCAGAGTACTGCGAGAT 371
 QY 751 GGGCGAGCTTTGAGAGAGGCGTGGCATCCGGCGGAGCCGAGCGCTGGCGCGGAGCTG 810
 DB 372 GGGCGAGCTTTGAGAGAGGCGTGGCATCCGGCGGAGCCGAGCGCTGGCGCGGAGCTG 431
 QY 811 GAGGTGTTGGGAGGAGCCACCCGAGTCCGCTCAAGGAGACCTGGGCTGTGTGTTGT 870
 DB 432 GAGGTGTTGGGAGGAGCCACCCGAGTCCGCTCAAGGAGACCTGGGCTGTGTGTTGT 491
 QY 871 GACATCAAGTCTCAGAGCTCTCTATCTGAGCGCTTGGGGGCGACTAAGTGAAGTGC 930
 DB 492 GACATCAAGTCTCAGAGCTCTCTATCTGAGCGCTTGGGGGCGACTAAGTGAAGTGC 551
 QY 931 GCCCTGCTGAGGAGCCCTGCGGGGCGTGTCTGACTGAGGCGCTGCGAGAGCTGTGGGC 990
 DB 552 GCCCTGCTGAGGAGCCCTGCGGGGCGTGTCTGACTGAGGCGCTGCGAGAGCTGTGGGC 611
 QY 991 CGGAGAGCGTTCGCGCTGCTGAGTGTGAGTGAAGCGTGAATAGAGCTGGCGCGCGC 1050
 DB 612 CGGAGAGCGTTCGCGCTGCTGAGTGTGAGTGAAGCGTGAATAGAGCTGGCGCGCGC 671
 QY 1051 CGCCTGTTGCTGA 1063
 DB 672 CGCCTGTTGCTGA 684

RESULT 17
 LOCUS BX384908/c 1027 bp mRNA linear EST 27-APR-2004
 DEFINITION BX384908 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
 Homo sapiens cDNA clone CSDBL0041YA21 3-PRIME, mRNA sequence.
 ACCESSION BX384908 GI:30455279
 VERSION BX384908
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1027)
 Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
 TITL Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seque@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with NotI and cloned
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 9074.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?b=CSDBL0041YA11NP1lec=9074.r.
 FEATURES
 source
 1..1027
 /organism="Homo sapiens"
 /mol_type="mRNA"

ORIGIN

/db xref="taxon:9606"
 /clone="CS0DL004YA21"
 /cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /cell_line="RAMOS CELL LINE"
 /clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 32.2%; Score 607; DB 5; Length 1027;
 Best Local Similarity 99.8%; Pred. No. 5.6e-298;
 Matches 657; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1177 CCCCTTGAGAGCCCTCCAGAGATGGGGCTGAGGCTTAACCTTCAGCTGAG 1236
 DB 674 CCCCTTGAGAGCCCTCCAGAGATGGGGCTGAGGCTTAACCTTCAGCTGAG 615
 QY 1237 TTTCCTTCCAGACTCTCTTACCCAGGTGTGCCCTTTAGCCTCCGAGGCGGGAGC 1296
 DB 614 TTTCCTTCCAGACTCTCTTACCCAGGTGTGCCCTTTAGCCTCCGAGGCGGGAGC 555
 QY 1297 TGGGCTGTATCTCAAGAGAGGGGACAGCTACACTCAACAAAGCCCTCTGCAC 1356
 DB 554 TGGGCTGTATCTCAAGAGAGGGGACAGCTACACTCAACAAAGCCCTCTGCAC 495
 QY 1357 ATTGTATCTGTATCTTGGGCTGTGTGACCTGTACAGAGGTCACACTGCTGATGCTC 1416
 DB 494 ATTGTATCTGTATCTTGGGCTGTGTGACCTGTACAGAGGTCACACTGCTGATGCTC 435
 QY 1417 ACACTGCCCCCTGCTGAGATCTTCCCTGGGCTTGCCTTGGCTTCCGACACACAC 1476
 DB 434 ACACTGCCCCCTGCTGAGATCTTCCCTGGGCTTGCCTTGGCTTCCGACACACAC 375
 QY 1477 TTCTTTGGCTTAAGGGCTTCTCTCTCAGAGACCTTAAATTGACCAACCAACTGGGCT 1536
 DB 374 TTCTTTGGCTTAAGGGCTTCTCTCTCAGAGACCTTAAATTGACCAACCAACTGGGCT 315
 QY 1537 TCAGCCACATCAATGGGCACTGAGAGCTGGGGGTGCACATGGGGGCTGCTCACTTGGCCAC 1596
 DB 314 TCAGCCACATCAATGGGCACTGAGAGCTGGGGGTGCACATGGGGGCTGCTCACTTGGCCAC 255
 QY 1597 ACATCTCCAGCCAGCAGAGGCGCTGCGCAGCTTCAATTACAGACCTGATCTCTCAAC 1656
 DB 254 ACATCTCCAGCCAGCAGAGGCGCTGCGCAGCTTCAATTACAGACCTGATCTCTCAAC 195
 QY 1657 TTCCCCCTGCTGTCCAGAGCTGAACATGACTTGGACTTGGATGTCACTTGGAGTGTCA 1716
 DB 194 TTCCCCCTGCTGTCCAGAGCTGAACATGACTTGGACTTGGATGTCACTTGGAGTGTCA 135
 QY 1717 CATGGAGTGTATGGAGCATATACCAAGGCTACTGTTGACATGGGGGCAAAACCA 1776
 DB 134 CATGGAGTGTATGGAGCATATACCAAGGCTACTGTTGACATGGGGGCAAAACCA 75
 QY 1777 GTAAACAGCCACTTCTTGGAAAAGGAATCAAAAGCTTTTGAGGATGAAAAGAC 1834
 DB 74 GTAAACAGCCACTTCTTGGAAAAGGAATCAAAAGCTTTTGAGGATGAAAAGAC 17

RESULT 18
 CK429257 655 bp mRNA linear EST 06-JAN-2004
 LOCUS o134b12.y1 Human lacrimal gland, unamplified: o134b12.y1
 DEFINITION clone o134b12 5', mRNA sequence.
 ACCESSION CK429257
 VERSION CK429257.1 GI:40676160
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 655)
 AUTHORS Dickinson,D., Laurie,G. and Wistow,G.
 TITLE Expressed sequence tag analysis of human lacrimal gland
 JOURNAL Unpublished (2002)
 COMMENT Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/31, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graham@helix.nih.gov
 Plate: 34 row: b column: 12
 Seq primer: M13RPI reverse primer (ABI).
 Location/Qualifiers

FEATURES
 source
 1..655
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="o134b12"
 /issue_type="lacrimal gland"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /clone_lib="Human lacrimal gland, unamplified: o1"
 /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted from 2 human lacrimal glands. A directionally cloned cDNA library in the pCMVSPORT6 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGAGAGGGGCGGCTT)5-3']. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 32.1%; Score 604; DB 7; Length 655;
 Best Local Similarity 99.8%; Pred. No. 1.9e-296;
 Matches 654; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 52 CCGTGCCCCAGGAGTCATCGAGCGCCAGATCTGAGCGGGTCTGAGCTTGTCCGCTC 111
 DB 1 CCGTGCCCCAGGAGTCATCGAGCGCCAGATCTGAGCGGGTCTGAGCTTGTCCGCTC 60
 QY 112 CCTCCCCCGGAAATGGCGCTATCCGGGTGACCCCGGCGCTGCTGGAGAGATGAG 171
 DB 61 CCTCCCCCGGAAATGGCGCTATCCGGGTGACCCCGGCGCTGCTGGAGAGATGAG 120
 QY 172 TGGCTGAGTACTTACGGGATGCTGTGCTTACCCGATTTTGAAGTGGCGGGCA 231
 DB 121 TGGCTGAGTACTTACGGGATGCTGTGCTTACCCGATTTTGAAGTGGCGGGCA 180
 QY 232 CTGACCGAGTGGAGCTGAGCTCTGAGCTTCTGCTGATGAGGCTCTGAGCGCGCC 291
 DB 181 CTGACCGAGTGGAGCTGAGCTCTGAGCTTCTGCTGATGAGGCTCTGAGCGCGCC 240
 QY 292 GGAAGCTTAAGCCCGGCGCCGAGCGGCTTGAAGCTTCTGAGAGTGGAGCGCGGG 351
 DB 241 GGAAGCTTAAGCCCGGCGCCGAGCGGCTTGAAGCTTCTGAGAGTGGAGCGCGGG 300
 QY 352 CAGTGGCGGAGAGCACTGGCGGCTGTGGGGCACTCTCGCGCTGTGGCGCGGCAC 411
 DB 301 CAGTGGCGGAGAGCACTGGCGGCTGTGGGGCACTCTCGCGCTGTGGCGCGGCAC 360
 QY 412 GACCTGCTGCGGACCTTGGCGGCGCAAGCGGCGCGGAGTGTCTCAGAACGCTATTAGC 471
 DB 361 GACCTGCTGCGGACCTTGGCGGCGCAAGCGGCGCGGAGTGTCTCAGAACGCTATTAGC 420
 QY 472 TATGGACCTCCAGCTCTTCAAGAGAGACAGAGGATGCTGCGCTGCGCGGAGTCA 531
 DB 421 TATGGACCTCCAGCTCTTCAAGAGAGACAGAGGATGCTGCGCTGCGCGGAGTCA 480

```

QY 532 AGCAGTTCTGCAGAAATTCCTGACAGAGGTCACTGGAGAGACAGGCTTCCCCCAACCAACGCG 591
DB 481 AGCAGTTCTGCAGAAATTCCTGACAGAGGTCACTGGAGAGACAGGCTTCCCCCAACCAACGCG 540
QY 592 CACGGCGGAGAGTGGGGGCGGGCCAGTGTGTGTCCAGAGGCGGGCGGAGAGAGGGGCCCCA 651
DB 541 CACGGCGGAGAGTGGGGGCGGGCCAGTGTGTGTCCAGAGGCGGGCGGAGAGAGGGGCCCCA 600
QY 652 GCCGCAACCCAGAGAGAGTCAAGAGCCCGCAGACCTTCTCTGAGAGCAAGTGA 706
DB 601 GCCGCAACCCAGAGAGAGTCAAGAGCCCGCAGACCTTCTCTGAGAGCAAGTGA 655

RESULT 19
BI907024 634 bp mRNA linear EST 16-OCT-2001
LOCUS 603064980F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214211 5'
DEFINITION mRNA sequence.
ACCESSION BI907024
VERSION BI907024.1 GI:16169804
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM1537 row: p column: 20
High quality sequence stop: 632.
Location/Qualifiers
1. 634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5214211"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/notes="Vector: PCMV-SPORE; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC library."

ORIGIN
Query Match 32.0%; Score 603; DB 4; Length 634;
Best Local Similarity 100.0%; Pred. No. 6e-296;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 451 GTGTCTCCAGAAAGCTATGCACTTCCAGCTCTTCAAGAGAGAGAGGAGTGC 510
DB 32 GTGTCTCCAGAAAGCTATGCACTTCCAGCTCTTCAAGAGAGAGAGGAGTGC 91
QY 511 TGGCGTGGCGGTGGAGAGTTCGCAAAATTCAGAGAGGGTCACTGGAGAGACA 570
DB 92 TGGCGTGGCGGTGGAGAGTTCGCAAAATTCAGAGAGGGTCACTGGAGAGACA 151
QY 571 GGCTCCCCCAACCAAGCGGAGCGGAGTGGGGCGGCCCAAGTGTGTGTCCAGA 630

```

```

DB 152 GGCTCCCCCAACCAAGCGGAGCGGAGTGGGGCGGCCCAAGTGTGTGTCCAGA 211
QY 631 CGGCGGAGAGAGGGGCGCCAGCGGAGCCGAGCAGACAGTCAAGGCCGAGACCTTCC 690
DB 212 CGGCGGAGAGAGGGGCGCCAGCGGAGCCGAGCAGACAGTCAAGGCCGAGACCTTCC 271
QY 691 TCTGAAGGCAAGTGAAGTGTGATCCGAGTCCGGGTTTCAAGCAGATGACTGAGCAT 750
DB 272 TCTGAAGGCAAGTGAAGTGTGATCCGAGTCCGGGTTTCAAGCAGATGACTGAGCAT 331
QY 751 GGGCCAGCCTTGGAGAGAGGCGTGGCATCCGGCGGCCAGGCGCTGGCGGAGCTG 810
DB 332 GGGCCAGCCTTGGAGAGAGGCGTGGCATCCGGCGGCCAGGCGCTGGCGGAGCTG 391
QY 811 GAGCTGTGGGAGAGGCAACCCGAGTGGCGCTCAAGGAGACCTGGGCTGTGTTGT 870
DB 392 GAGCTGTGGGAGAGGCAACCCGAGTGGCGCTCAAGGAGACCTGGGCTGTGTTGT 451
QY 871 GACATCAAGTCTCAGAGCTCTCTATCTGACGCGCTTGGGGCAGTACCTGATGAGC 930
DB 452 GACATCAAGTCTCAGAGCTCTCTATCTGACGCGCTTGGGGCAGTACCTGATGAGC 511
QY 931 GCCCTCTGAGGCGCTTGGCGGGGCTGTTCTGACTGAGGCGCTGCAAGGCTGTGGGC 990
DB 512 GCCCTCTGAGGCGCTTGGCGGGGCTGTTCTGACTGAGGCGCTGCAAGGCTGTGGGC 571
QY 991 CGGAGAGCTGTTCCGCTGTGCTGATGTGATGAGAGCTGACATAGAGTGGCGCGGC 1050
DB 572 CGGAGAGCTGTTCCGCTGTGCTGATGTGATGAGAGCTGACATAGAGTGGCGCGGC 631
QY 1051 CGC 1053
DB 632 CGC 634

RESULT 20
B0956773 991 bp mRNA linear EST 21-AUG-2002
LOCUS AGENCOURT_8804443 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6377779
DEFINITION 5' mRNA sequence.
ACCESSION B0956773
VERSION B0956773.1 GI:22372251
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLCM2562 row: b column: 20
High quality sequence stop: 542.
Location/Qualifiers
1. 991
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6377779"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_47"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:

```


[illegible]

RESULT 22	LOCUS	DEFINITION	BM670485/c
618 bp	linear	EST 27-FEB-2002	BM670485
UI-E-DWI-aha-1-04-0-UI s1	UI-E-DWI Homo sapiens cDNA clone		
UI-E-DWI-aha-1-04-0-UI 3'	mRNA sequence.		

VERSION	BM670485.1	GI:18980
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 618)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Donald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reegen.com).
Seq primer: M13 Forward
SOLVA=yes.

FEATURES	Location/Qualifiers
source	1. .618

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-DWI-aha-1-04-0-UI"
 /tissue_type="lens"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_1ib="UI-E-DWI"
 /note="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a
 modified polylinker. Site_1: EcoR I, Site_2: Not I;
 UI-E-DWI is a normalized cDNA library containing the
 following tissue(s): lens. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pRTT3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CAGTTGCGA. This library
 was created for the program, Gene Discovery in the Visual
 System, supported by National Eye Institute (NEI).
 TAG_TISSUE=human lens

```

TAG_LIB=UI-E-DW1
TAG_SEQ=CGATTAGCGA'
ORIGIN

```

Query Match	31.8%;	Score 599;	DB 4;	Length 618;
Best Local Similarity	100.0%;	Pred. No. 6.6e-294;		
Matches 599;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1285	GGAGGCGGGGGCTGGGGCTGTATCTCGAAGGGGGGGGACAGCTACACATCCACCAAG	1344
Dp	618	GGAGGCGGGGGCTGGGGCTGTATCTCGAAGGGGGGGGACAGCTACACATCCACCAAG	559
Qy	1345	GCCCCCTGCACATTTGTATCTGTATCTTGGGCTGTGTCACTGTCA CAGGTGCA CAC	1404
Dp	558	GCCCCCTGCACATTTGTATCTGTATCTTGGGCTGTGTCACTGTCA CAGGTGCA CAC	499
Qy	1405	TGCGTATGTCTACACCTGCCCCCTGTGAGATCTTCCCTGGGCTCTGCCCCCTGCTT	1464
Dp	498	TGCGTATGTCTACACCTGCCCCCTGTGAGATCTTCCCTGGGCTCTGCCCCCTGCTT	439
Qy	1465	CCGAGCACACCTCTTTGGCTTAAAGGCTTCTCTCAGGACCTCTTAATTGACACAA	152
Dp	438	CCGAGCACACCTCTTTGGCTTAAAGGCTTCTCTCAGGACCTCTTAATTGACACAA	379
Qy	1525	CCAACCTGGGCTTCAGGCCACATCAGTGGGCACTGGAGCTGGGGTGCATATGGGGCTGCT	1584
Dp	378	CCAACCTGGGCTTCAGGCCACATCAGTGGGCACTGGAGCTGGGGTGCATATGGGGCTGCT	319
Qy	1585	CACCTTGGCCACACATCTCCAGGCCAGCGAGGGCCCTGCCCCAGCTTCAATTTACAGACTG	1644
Dp	318	CACCTTGGCCACACATCTCCAGGCCAGCGAGGGCCCTGCCCCAGCTTCAATTTACAGACTG	259
Qy	1645	ACTCTCTCACTTCCCCCGTGTCCAGAGCTGGAACATAGACTTTCACCTTGGATGTCA	1704
Dp	258	ACTCTCTCACTTCCCCCGTGTGTCCAGAGCTGGAACATAGACTTTCACCTTGGATGTCA	199
Qy	1705	CCTGAGTGTCACTGGAGTGTATTGGCAGCATCATACCAAGGCTTACTGTGTCA CATTG	1764
Dp	198	CCTGAGTGTCACTGGAGTGTATTGGCAGCATCATACCAAGGCTTACTGTGTCA CATTG	139
Qy	1765	GGGCGAAAACCACTAAACAGCCACCTTCTTGGAAAAGGGAATGCAAGGCTTTGGGGGTGA	1824
Dp	138	GGGCGAAAACCACTAAACAGCCACCTTCTTGGAAAAGGGAATGCAAGGCTTTGGGGGTGA	79
Qy	1825	TGGAAAAGACCTTTTAAATGATACCAATTTAACTGCCCCGTGAAAAGGGCATAGGTGGG	1883
Dp	78	TGGAAAAGACCTTTTAAATGATACCAATTTAACTGCCCCGTGAAAAGGGCATAGGTGGG	20

RESULT 23	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BM928759	1028 bp	mRNA	linear	EST 12-MAR-2002								
AGENCOURT_6726912	NIH_MGC_100	Homo sapiens cDNA clone IMAGE:5798907	5', mRNA sequence.									
BM928759												
BM928759.1	GI:19379138	EST.										
							Homo sapiens (human)					
							Homo sapiens					
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
							Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
							1 (bases 1 to 1028)					
							NIH-MGC http://mgs.nci.nih.gov/ .					
							National Institutes of Health, Mammalian Gene Collection (MGC)					
							Unpublished (1999)					
							Contact: Robert Strausberg, Ph.D.					
							Email: cgap@femail.nih.gov					
							Tissue Procurement: CGAP (Stanford)					
							cDNA Library Preparation: Rubin Laboratory					
							cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)					
							DNA Sequencing by: Agencourt Bioscience Corporation					
							Clone distribution: MGC clone distribution information can be					
							found through the I.M.A.G.E. Consortium/LNL at:					

http://image.llnl.gov
 Plate: L1CM2027 row: k column: 04
 High quality sequence stop: 585.
 Location/Qualifiers

FEATURES

source

```

1. 1028
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5798907"
/issue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_100"
/note="Organ: liver; Vector: pOTB7, Site_1: XhoI, Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NH_MGC
Library."

```

ORIGIN

Query Match 31.8%; Score 599; DB 5; Length 1028;
 Best Local Similarity 100.0%; Pred. No. 6.8e-294;
 Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 583 ACCAAGCGGAGGCGGAGTCCGGGCGGCGGCGGAGTGTGTGTCAGACGCGCGGAGAGA 642
DB 1 ACCAAGCGGAGGCGGAGTCCGGGCGGCGGCGGCGGAGTGTGTGTCAGACGCGCGGAGAGA 60
OY 643 GGGGCGCCAGCGCGACCCCGACGAGTACAGAGCCCGGACCTTCTGTGAAGGCAAA 702
DB 61 GGGGCGCCAGCGCGACCCCGACGAGTACAGAGCCCGGACCTTCTGTGAAGGCAAA 120
OY 703 GTGACCTGTGACATCCGGCTCCGGGTTCCAGCAGATCTGCGAGATGGGCGGAGCTTG 762
DB 121 GTGACCTGTGACATCCGGCTCCGGGTTCCAGCAGATCTGCGAGATGGGCGGAGCTTG 180
OY 763 GAGCAGGCGGCGGAGTCCGGGCGGCGGCGGCGGAGTGTGTGTCAGACGCTGTGAG 822
DB 181 GAGCAGGCGGCGGAGTCCGGGCGGCGGCGGCGGAGTGTGTGTCAGACGCTGTGAG 240
OY 823 CAGGCGCAGCGAGTGTGCGCTCAAGGGACTGTGTGTGTGTGTGATCAAGATTG 882
DB 241 CAGGCGCAGCGAGTGTGCGCTCAAGGGACTGTGTGTGTGTGTGATCAAGATTG 300
OY 883 TCAGACTCTCTCATCTGACAGCGCTTGTGGGCGGAGTACTGTGTGTGTGTGTGAG 942
DB 301 TCAGACTCTCTCATCTGACAGCGCTTGTGGGCGGAGTACTGTGTGTGTGTGTGAG 360
OY 943 GCGCTCGGGGCGGCTGTTCTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTT 1002
DB 361 GCGCTCGGGGCGGCTGTTCTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTT 420
OY 1003 CCGCTCTGCTGCTGATGTGATGAGGCTGATATAGGCTGTGGGCGGCGGCTGTGCTG 1062
DB 421 CCGCTCTGCTGCTGATGTGATGAGGCTGATATAGGCTGTGGGCGGCGGCTGTGCTG 480
OY 1063 ATGAGAGAGAGAGGGGGGGGGCGGCGGCGGAGAGAGGCTTCTGATCCAGAGTGGCAGATT 1122
DB 481 ATGAGAGAGAGAGGGGGGGGGCGGCGGCGGAGAGAGGCTTCTGATCCAGAGTGGCAGATT 540
OY 1123 GATCCCACTCCAAAGTCTCCGGGCGACCTTCTCTGTGGAGAGAGCAGCATCTTACCCCT 1181
DB 541 GATCCCACTCCAAAGTCTCCGGGCGACCTTCTCTGTGGAGAGAGCAGCATCTTACCCCT 599

```

RESULT 24
 CB321808/c 732 bp mRNA linear EST 05-MAR-2003
 LOCUS
 DEFINITION UI-CF-EN1-aej-j-24-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
 UI-CF-EN1-aej-j-24-0-UI 3', mRNA sequence.

ACCESSION CB321808
 VERSION CB321808.1 GI:28856466
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Donald, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.research.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

```

1. 732
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-aej-j-24-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTGCTCAGGT.
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_LIB=UI-CF-EN1
TAG_SEQ=CTGCTCAGGT"

```

ORIGIN

Query Match 31.8%; Score 598; DB 6; Length 732;
 Best Local Similarity 99.7%; Pred. No. 2.2e-293;
 Matches 698; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1177 CCCCTTGACAGCCCTCCACAGATGTGGCTGTGAGGCTTAACCATTTCCAGCTGAG 1236
DB 725 CCCCTTGACAGCCCTCCACAGATGTGGCTGTGAGGCTTAACCATTTCCAGCTGAG 666
OY 1237 TTTCCTTCCAGACTCTCTCAATCCCGGAGTGTGCCCCCTTATGACTCCGAGGCGGAGGC 1296
DB 665 TTTCCTTCCAGACTCTCTCAATCCCGGAGTGTGCCCCCTTATGACTCCGAGGCGGAGGC 606

```


QY	1297	TGGGCGCTGATCTCAGAAAGGGAGGGGACACTGACATCACTCAACCAAGAGCCCTCTGAC	1356
Db	605	TGGGCGCTGATCTCAGAAAGGGAGGGGACAGANTACACTCAACCAAGAGCCCTCTGAC	546
QY	1357	ATTGATATCTGATCTTGGGCTGTCTGTGACATGTGCACAGTGCACACTCGCTCATGCTC	1416
Db	545	ATTGATATCTGATCTTGGGCTGTCTGTGACATGTGCACAGTGCACACTCGCTCATGCTC	486
QY	1417	ACACTGCCCCCTGCTGAGATCTTCCCTGGGCTCTGCCCCCTGGCTCTTCCAGACACAC	1476
Db	485	ACACTGCCCCCTGCTGAGATCTTCCCTGGGCTCTGCCCCCTGGCTCTTCCAGACACAC	426
QY	1477	TTCTTTGGCCCTAAGGGCTCTCTCTGCAGGACCTCTAATTTCACACAAACCAACTGGGCT	1536
Db	425	TTCTTTGGCCCTAAGGGCTCTCTCTGCAGGACCTCTAATTTCACACAAACCAACTGGGCT	366
QY	1537	TCAGCCACATCATGTGGGACCTGTGAGCTGGGGTGCACATGGGGCTGTCTCATCTTGCCAC	1596
Db	365	TCAGCCACATCATGTGGGACCTGTGAGCTGGGGTGCACATGGGGCTGTCTCATCTTGCCAC	306
QY	1597	ACATCTCCAGCCAGCCAGGCGCCCTGCCCCAGCTTCAATTATACAGACTGTCTCTCAAC	1656
Db	305	ACATCTCCAGCCAGCCAGGCGCCCTGCCCCAGCTTCAATTATACAGACTGTCTCTCAAC	246
QY	1657	TTCCCCCTGTGTGTCAGAGCTGAAACAATGACCTTGACCTTGGAGTGCACCTGGAGTGTCA	1716
Db	245	TTCCCCCTGTGTGTCAGAGCTGAAACAATGACCTTGACCTTGGAGTGCACCTGGAGTGTCA	186
QY	1717	CATGGAGATGTTATGGCGAGCATCATACCAAGGCTACTGTGGACATGGGGCCAAAACA	1776
Db	185	CATGGAGATGTTATGGCGAGCATCATACCAAGGCTACTGTGGACATGGGGCCAAAACA	126
QY	1777	GTAACAAGCCACCTTCTTGGAAAGGGAATGCAGAGGCTTTGGGGGTGATGGAAGAAGCT	1836
Db	125	GTAACAAGCCACCTTCTTGGAAAGGGAATGCAGAGGCTTTGGGGGTGATGGAAGAAGCT	66
QY	1837	TTTCAAAATGATATACCAATTAACCTGCTCTGAAAGGGCAT	1876
Db	65	TTTCAAAATGATATCAATTAACCTGCTCTGAAAGGGCAT	26
RESULT 25			
LOCUS	BE907425		
DEFINITION	BE907425	1131 bp	mRNA linear EST 20-OCT-2000
ACCESSION	601500161P1	NIH_MGC_70	Homo sapiens cDNA clone IMAGE:390194 5',
VERSION	BE907425		mRNA sequence.
KEYWORDS	BE907425.1	GI:10400971	
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Mumayyoti, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Eukaryota; Euteleostomi; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 1131)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strusberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)		
	Genome Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLNL at:		
	http://image.lnl.gov		
	Plate: LHAM9703 row: 1 column: 15		
	High quality sequence stop: 686.		
FEATURES			
source			
	1..1131		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	Location/Qualifiers		

ORIGIN	Query Match	31.5%; Score 594; DB 2; Length 1131;
	Best Local Similarity 100.0%; Pred. No. 2,4e-291;	
	Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	451 GTGTTCTCAGAACGCTATAGCTATAGCACTTCACGCTTTCAAGAGAGACAGAGGTAGC 510	
Db	30 GTGCTTCAGAACGCTATAGCTATAGCACTTCACGCTTTCAAGAGAGACAGAGGTAGC 89	
QY	511 TGCCTGCGCGTCGCGAGCTCAAGAGTTTCGCAATTCTCAGCAGGGCTCAGTGGAGACA 570	
Db	90 TGCCTGCGCGTCGCGAGCTCAAGAGTTTCGCAATTCTCAGCAGGGCTCAGTGGAGACA 149	
QY	571 GGCCTCCCCCAACCAAGCGGACGGCGAGAGTGGGGGCGGCCAGTGTGTGTCCAGA 630	
Db	150 GGCCTCCCCCAACCAAGCGGACGGCGAGAGTGGGGGCGGCCAGTGTGTGTCCAGA 209	
QY	631 CGGCGGCGAGAGGGGGCCCAAGCCGACCCAGAGCAGTCAAGAGCCGCGACACTTCC 690	
Db	210 CGGCGGCGAGAGGGGGCCCAAGCCGACCCAGAGCAGTCAAGAGCCGCGACACTTCC 269	
QY	691 TCTGAAGGCAAGTGAAGCTGTGAATCCGGCTCCGGGTTGAGAGAGATCTGGCAGAT 750	
Db	270 TCTGAAGGCAAGTGAAGCTGTGAATCCGGCTCCGGGTTGAGAGAGATCTGGCAGAT 329	
QY	751 GGGCGAGCTTTGAGAGAGGGGGTGGCAATCCCGGGGCCCCAGAGGCGTGGCGGACGCTG 810	
Db	330 GGGCGAGCTTTGAGAGAGGGGGTGGCAATCCCGGGGCCCCAGAGGCGTGGCGGACGCTG 389	
QY	811 GACGTGTTTGGGACAGGCCACCGCAGTGTGGCTCAAGAGGACTGTGGCTGTGTGTTGT 870	
Db	390 GACGTGTTTGGGACAGGCCACCGCAGTGTGGCTCAAGAGGACTGTGGCTGTGTGTTGT 449	
QY	871 GACATCAAGTTCTCAGAGCTTCTCTATCTGGAAGCCTTTTGGGGCGACTTACTGAGTGGC 930	
Db	450 GACATCAAGTTCTCAGAGCTTCTCTATCTGGAAGCCTTTTGGGGCGACTTACTGAGTGGC 509	
QY	931 GCCCTGCTGACAGGCGCTCGCGGGGGGCTGTCTCTGCTGAGGCGCTGCGAGAGGCTGTGGC 990	
Db	510 GCCCTGCTGACAGGCGCTCGCGGGGGGCTGTCTCTGCTGAGGCGCTGCGAGAGGCTGTGGC 569	
QY	991 CGGAGGCGCTGTCCCTGCTGTGTGTGAGTGTGAGTGAAGGCTGACTTGAAGGCTGAC 1044	
Db	570 CGGAGGCGCTGTCCCTGCTGTGTGTGAGTGTGAGTGAAGGCTGAAGGCTGAC 623	
RESULT 26	BG251151 927 bp mRNA linear EST 13-FEB-2001	
LOCUS	BG251151	
DEFINITION	60236695AF1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473349 5',	
ACCESSION	NRNA sequence.	
KEYWORDS	BG251151	
SOURCE	BG251151.1 GI:12760967	
ORGANISM	Homo sapiens (human)	
REFERENCE	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
TITLE	NIH-MGC http://mgc.nci.nih.gov/.	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999)	
	Contact: Robert Strausberg, Ph.D.	
	Email:cgapds-remail.nih.gov	

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9abs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: L1AM12771 row: a column: 10
 High quality sequence stop: 651.
 Location/Qualifiers
 1..1051
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5746449"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_120"
 /note="Organ: pooled pancreas and spleen; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of spleen and pancreas from 28 yo
 male. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert
 size 1.5 Kb, insert size range 1-2.5 Kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 025. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 30.6%; Score 576; DB 5; Length 1051;
 Best Local Similarity 99.8%; Pred. No. 3.8e-282;
 Matches 626; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 451 GTGCTCCAGAGGCTAGCTAGTACCTCCAGCTTCAAGAGGAGAGGGTAC 510
 DB 62 GTGCTCCAGAGGCTAGCTAGTACCTCCAGCTTCAAGAGGAGAGGGTAC 121
 QY 511 TGGCGTGGCGGTGGAGTCAAGAGTTCGAAATTCAGAGGGGTCAAGAGACA 570
 DB 122 TGGCGTGGCGGTGGAGTCAAGAGTTCGAAATTCAGAGGGGTCAAGAGACA 181
 QY 571 GGCCTCCCCCAACCAAGCGGACGGCGGAGTGGGGCCGGCCAGTGTGTCACA 630
 DB 182 GGCCTGGCGCCCAACCAAGCGGACGGCGGAGTGGGGCCGGCCAGTGTGTCACA 241
 QY 631 CGGCGGCGAGAGGGGCGGACCGGACCGAGAGGTCAGAGCCCGCAGACCTTCC 690
 DB 242 CGGCGGCGAGAGGGGCGGACCGGACCGAGAGGTCAGAGCCCGCAGACCTTCC 301
 QY 691 TCTGAAGGCAAGTGAATCTGTGATCCGGCTCCGGGTTGAGACAGATCTCGAGCAT 750
 DB 302 TCTGAAGGCAAGTGAATCTGTGATCCGGCTCCGGGTTGAGACAGATCTCGAGCAT 361
 QY 751 GGGCGAGCCTTGGAGACGGGCGTGGCATCCCGGGCGCCCAAGGGGCTGGCGGACGCTG 810
 DB 362 GGGCGAGCCTTGGAGACGGGCGTGGCATCCCGGGCGCCCAAGGGGCTGGCGGACGCTG 421
 QY 811 GAGGTGTTGGGAGAGGCGGACCGAGTGTCCGCTCAAGGGGACCTGGGCTCTGTGGTTGT 870
 DB 422 GAGGTGTTGGGAGAGGCGGACCGAGTGTCCGCTCAAGGGGACCTGGGCTCTGTGGTTGT 481
 QY 871 GACATCAAGTTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGGCGACTACTGAGTGGC 930
 DB 482 GACATCAAGTTCAGAGCTCTCTATCTGAGAGCCTTCTGGGGGCGACTACTGAGTGGC 541
 QY 931 GCCCTGCTGAGGCGCTGCGGGGCGTGTCTTGACTGAGGCGCTTCTGAGAGGCTGTGGGC 990
 DB 542 GCCCTGCTGAGGCGCTGCGGGGCGTGTCTTGACTGAGGCGCTTCTGAGAGGCTGTGGGC 601
 QY 991 CGGAGGCTTTGGCGCTGTGTCAGTGTGATGAGGCTGATATAGAGGCTTGGCCGGCGC 1050

DB 602 CGGAGGCTGTTCCGCTGTGTGTCAGTGTGATGAGGCTGATGAGGCTGGCCGGCGC 661
 QY 1051 CGGCTGTTCTGATGAGGAGGAGG 1077
 DB 662 CGGCTGTTCTGATGAGGAGGAGG 688
 RESULT 30
 CR819031/c 571 bp mRNA linear EST 11-MAR-2004
 LOCUS 1506C09.x5 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 DEFINITION cDNA clone IMAGE:5675536 3', mRNA sequence.
 ACCESSION CR819031
 VERSION CR819031.1 GI:44835956
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 571)
 Melton, D., Meadows, A., Clifton, S., Hillier, L., Marra, M., Pape, D.,
 Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B.,
 Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M.,
 McCann, R., Cole, R., Tsagaris, R., Williams, T., Jackson, Y. and
 Bowers, Y.
 Washu-Harvard Pancreas EST Project
 Unpublished (2000)
 CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Other ESTs: 1506C09.y1
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 This read is a 3' RESEQUENCE of a previously sequenced pancreas
 clone
 This read has been verified (found to hit its original self in the
 correct orientation)
 Seq primer: -40UP from Gibco.
 Location/Qualifiers
 1..571
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5675536"
 /sex="Both"
 /tissue_type="Islets of Langerhans"
 /dev_stage="Adult"
 /lab_host="DH10B"
 /clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
 /note="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
 Site 2: Sal 1; Starting library constructed using
 Superscript Plasmid library kit (Life Technologies). cDNA
 made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 Kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."
 ORIGIN
 Query Match 30.2%; Score 569; DB 7; Length 571;
 Best Local Similarity 100.0%; Pred. No. 1.4e-278;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1315 GGGAGGGGACACGACTACACCAAGGCCCCCTGCACTGTATCTGATCTTGG 1374


```

ACCESSION  CN373290
VERSION      CN373290.1
KEYWORDS    GI:47373224
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       1 (bases 1 to 644)
JOURNAL     Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraige, J., Fisk, G.J.,
COMMENT     Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
            Lebkowski, J and Stanton, L.W.
            Transcriptome characterization elucidates signaling networks that
            control human ES cell growth and differentiation
            Nat. Biotechnol. 22 (6), 707-716 (2004)
            Contact: Brandenberger R
            Regenerative Medicine
            Geron Corporation
            230 Constitution Drive, Menlo Park, CA 94025, USA
            Tel: 650 473 8658
            Fax: 650 473 7760
            Email: rbrandenberger@geron.com
            Insert Length: 644 Std Error: 0.00.
            Location/Qualifiers
                1..644
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /tissue_type="embryonic stem cells, cell lines H1, H7, and
                H9"
                /note="To oligo dT primed, full-length enriched cDNA library
                from undifferentiated hES cell lines H1 (p32), H7 (p29),
                and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match      30.0%; Score 564; DB 7; Length 644;
Best Local Similarity 99.8%; Pred. No. 5e-276; Indels 0; Gaps 0;
Matches 614; Conservative 0; Mismatches 1;

QY 451 GTGTCTCCAGAAAGCTATAGCTATGACACCTCCAGCTCTTCAAGAGAGACAGAGGATAGC 510
DB 30 GTGTCTCCAGAAAGCTATAGCTATGACACCTCCAGCTCTTCAAGAGAGAGAGGATAGC 89
QY 511 TGGCGTGGCGGTGGCGAGTCAGACAGTCTTGCAGAAATTTCTCAGCAGGAGTCAGTGGAGACA 570
DB 90 TGGCGTGGCGGTGGCGAGTCAGACAGTCTTGCAGAAATTTCTCAGCAGGAGTCAGTGGAGACA 149
QY 571 GGGTCCCCCCCAACGAGCGGCGGAGTCGGGGCCGGCCCAAGTGGTGGTGGCAGA 630
DB 150 GGGTCCCCCCCAACGAGCGGCGGAGTCGGGGCCGGCCCAAGTGGTGGTGGCAGA 209
QY 631 CGGCGGCGAGAGAGGGGCCCGGACCGGACCCAGCAGAGTCAGAGCCCGGACAGCTTCC 690
DB 210 CGGCGGCGAGAGAGGGGCCCGGACCGGACCCAGCAGAGTCAGAGCCCGGACAGCTTCC 269
QY 691 TCTGAAGGCAAGTGAACCTTGAACATCCGGCTCCGGGTTGAGACAAGTACTCCGAGCAT 750
DB 270 TCTGAAGGCAAGTGAACCTTGAACATCCGGCTCCGGGTTGAGACAAGTACTCCGAGCAT 329
QY 751 GGGCGAGCCTTGGAGAGAGGGCGTGGCATCCGGCGGCGCCAGGCGGTGGGCGGCGAGCTG 810
DB 330 GGGCGAGCCTTGGAGAGAGGGCGTGGCATCCGGCGGCGCCAGGCGGTGGGCGGCGAGCTG 389
QY 811 GAGCTGTTTGGAGAGCCGACCGAGTGTGCGCTCAAGAGGACCTGGGCTTGTGTTTGT 870
DB 390 GAGCTGTTTGGAGAGCCGACCGAGTGTGCGCTCAAGAGGACCTGGGCTTGTGTTTGT 449
QY 871 GACATCAAGTTCTCAAGAGCTCTCTATCTGAGAGCCTTCTGGGCGCACTAGAGTGC 930
DB 450 GACATCAAGTTCTCAAGAGCTCTCTATCTGAGAGCCTTCTGGGCGCACTAGAGTGC 509
QY 931 GCCCTGTGAGAGCCCTGCGGGCGGTCTCTGACTGAGGCGCTTGGAGAGGCTGTGGGC 990

```

```

DB 510 GCCCTGTGAGAGCCCTGCGGGCGGTCTCTGACTGAGGCGCCCTGCGAGAGGCTGTGGGC 569
QY 991 CGGAGAGCTGTGCGCTGTGTGTCAGTGTGATGAGGCTGACTATGAGGCTGGCGGCC 1050
DB 570 CGGAGAGCTGTGCGCTGTGTGTCAGTGTGATGAGGCTGACTATGAGGCTGGCGGCC 629
QY 1051 CGCCTGTGTGTGATG 1065
DB 630 CGCCTGTGTGTGATG 644

RESULT 33
B1117988
LOCUS      B1117988
DEFINITION 602867244P1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:5015843 5',
            mRNA sequence.
ACCESSION  B1117988
KEYWORDS   B1117988
SOURCE     B1117988.1 GI:14568889
ORGANISM   Homo sapiens (human)
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       1 (bases 1 to 718)
JOURNAL     NIH-MGC http://mgs.nci.nih.gov/
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-r@mail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: L16M1825 row: 9 column: 12
            High quality sequence stop: 717.
            Location/Qualifiers
                1..718
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5015843"
                /tissue_type="small cell carcinoma"
                /cell_line="MGC3"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH MGC 7"
                /note="Organ: lung; Vector: pORF7; Site 1: XhoI; Site 2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCAAGAG(G). Size-selected >500bp for average
                insert size 1.8kb. Library constructed by Ling Hong in
                the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

Query Match      30.0%; Score 564; DB 4; Length 718;
Best Local Similarity 99.6%; Pred. No. 5e-276; Indels 0; Gaps 0;
Matches 714; Conservative 0; Mismatches 3;

QY 36 AAACAAGTGGCTGGCGCGCTGGCCCAAGAGTCAATGGAGGCCAGAAATCTGGCGGGTTCT 95
DB 2 AAACAAGTGGCTGGCGCGCTGGCCCAAGAGTCAATGGAGGCCAGAAATCTGGCGGGTTCT 61
QY 96 GAGCTTGTTCGCGCTCCCTCCCGCGGGAATGGCGCTATCGGGTGCAGCCCGGCGCGTG 155
DB 62 GAGCTTGTTCGCGCTCCCTCCCGGGAATGGCGCTATCGGGTGCAGCCCGGCGCGTG 121
QY 156 CTGGAGAGAGATGAGTGTGCTGAGTCACTACCGAGATGTGTGCTTCAACCGATGTTCA 215
DB 122 CTGGAGAGAGATGAGTGTGCTGAGTCACTACCGAGATGTGTGCTTCAACCGATGTTCA 181

```

QY 216 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 275
|
|
|
Db 182 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 241
|
|
|
QY 276 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 335
|
|
|
Db 242 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 301
|
|
|
QY 336 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 395
|
|
|
Db 302 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 361
|
|
|
QY 396 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 455
|
|
|
Db 362 GGTGGTGGGGGCACTGACCGAGTGCAGAGCTCTCTGCTTCTGTGATGA 421
|
|
|
QY 456 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 515
|
|
|
Db 422 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 481
|
|
|
QY 516 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 575
|
|
|
Db 482 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 541
|
|
|
QY 576 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 635
|
|
|
Db 542 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 601
|
|
|
QY 636 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 695
|
|
|
Db 602 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 661
|
|
|
QY 696 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 752
|
|
|
Db 662 TCGAAGACGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCTATAGCT 718
|
|
|

RESULT 34
BU688308/c 638 bp mRNA linear EST 07-OCT-2002
LOCUS
DEFINITION
UI-CF-ECL-aea-o-13-0-UI.81 UI-CF-ECL Homo sapiens cDNA clone
UI-CF-ECL-aea-o-13-0-UI.3', mRNA sequence.
ACCESSION
BU688308
VERSION
BU688308.1 GI:23544974
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
COMMENT
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul.mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.research.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD
POLY-A-Tails
Location/Qualifiers

FEATURES

source
1. .638
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-ECL-aea-o-13-0-UI"
/tissue="lung"
/dev_stage="Adult and Fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pTZ19-Pac (Pharmacia) with a
modified polylinker; Site: 1: Ecor I; Site 2: Not I;
UI-CF-ECL is a normalized cDNA library containing the
following tissue(s): Normal lung from adult and from fetal
day 64, day 87, week 19 and week 42. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Ecor I
adaptor, digested with Not I, and cloned directionally
into pTZ19-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AACTGCTTAC.
TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
and 380-383
TAG_LIB=UI-CF-ECL
TAG_SEQ=AACTGCTTAC"

ORIGIN

Query Match 29.7%; Score 559; DB 5; Length 638;
Best Local Similarity 99.8%; Pred. No. 1.8e-273;
Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1264 AGGTGGCCCTTACCTTCGAGAGGCGGGGCTGGGCTGTATCTCAGAGGAGGGGC 1323
|
|
|
Db 638 AGGTGGCCCTTACCTTCGAGAGGCGGGGCTGGGCTGTATCTCAGAGGAGGGGC 579
|
|
|
QY 1324 ACAGCTACACACACACCAAGGCGCCCTGACATGTATCTGTATCTTGGGCTGTG 1383
|
|
|
Db 578 ACAGCTACACACACCAAGGCGCCCTGACATGTATCTGTATCTTGGGCTGTG 519
|
|
|
QY 1384 CACTGTACAGGTGACACACACCTGCTCATGCTCAGACCTGCTGAGATCTTCCCTG 1443
|
|
|
Db 518 CACTGTACAGGTGACACACACCTGCTCATGCTCAGACCTGCTGAGATCTTCCCTG 459
|
|
|
QY 1444 GGCCTTGGCCCTGCTTCCAGACACACCTTCTTGGCTTAAAGGCTTCTCTCA 1503
|
|
|
Db 458 GGCCTTGGCCCTGCTTCCAGACACACCTTCTTGGCTTAAAGGCTTCTCTCA 399
|
|
|
QY 1504 GGAACCTTAATTTGACCAACCAACCTGGCTTACGCCACATCACTGAGCTTGAAGCT 1563
|
|
|
Db 398 GGAACCTTAATTTGACCAACCAACCTGGCTTACGCCACATCACTGAGCTTGAAGCT 339
|
|
|
QY 1564 GGGGTGACATGGGGGCTGCTCACTTGGCCACACATCTCCAGCCAGGCGCTTGGC 1623
|
|
|
Db 338 GGGGTGACATGGGGGCTGCTCACTTGGCCACACATCTCCAGCCAGGCGCTTGGC 279
|
|
|
QY 1624 CAGCTTAATTTACAGACCTGACTCTCTCACTTCCCTCTGCTGTCCAGACTGAACA 1683
|
|
|
Db 278 CAGCTTAATTTACAGACCTGACTCTCTCACTTCCCTCTGCTGTCCAGACTGAACA 219
|
|
|
QY 1684 TACACTTGAACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1743
|
|
|
Db 218 TACACTTGAACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 159
|
|
|
QY 1744 CAAAGCTTACTGTGACATGGGGCCAAACAGTAAGACCACTTCTTGGAAAGGGA 1803
|
|
|
Db 158 CAAAGCTTACTGTGACATGGGGCCAAACAGTAAGACCACTTCTTGGAAAGGGA 99
|
|
|
QY 1804 ATGCAAAAGCTTTGGGGGTGATGAAAAGACCTTTTACAAATGATACCAATTAACTGCC 1863
|
|
|
Db 98 ATGCAAAAGCTTTGGGGGTGATGAAAAGACCTTTTACAAATGATACCAATTAACTGCC 39
|
|
|

QY 1864 CTGAAAAGGG 1873
 DB 38 CTGAAAAGGG 29

RESULT 35
 BU191631 918 bp mRNA linear EST 04-SEP-2002
 AGENCOURT 8045541 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6086350
 LOCUS 5', mRNA sequence.

ACCESSION BU191631
 VERSION BU191631.1 GI:22705602
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 918)
 NIH-MGC http://mgi.mgi.nhl.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: LNCM2321 row: C column: 23
 High quality sequence stop: 670.
 Location/Qualifiers
 1..918
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6086350"
 /tissue_type="epidermoid carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_102"
 /note="Organ: salivary gland; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 29.6%; Score 558; DB 5; Length 918;
 Best Local Similarity 100.0%; Pred. No. 5.9e-273;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1177 CCCTTGACAGCCCTCCACAGAGATGTGGCTCTGAGGCTTAACATTTCCAGCTGAG 1236
 DB 161 CCCTTGACAGCCCTCCACAGAGATGTGGCTCTGAGGCTTAACATTTCCAGCTGAG 220

QY 1237 TTTCTTCCCAAGCTCTCTTACCCCGAGGTGTGCCCCCTTAACTCCGAGGCGGGGC 1296
 DB 221 TTTCTTCCCAAGCTCTCTTACCCCGAGGTGTGCCCCCTTAACTCCGAGGCGGGGC 280

QY 1297 TGGGCGCTGTATCTCAAGAGGAGGGGACAGCTACACCTCAAGAGGCGGGGCCTGAC 1356
 DB 281 TGGGCGCTGTATCTCAAGAGGAGGGGACAGCTACACCTCAAGAGGCGGGGCCTGAC 340

QY 1357 ATTGATCTCTGATCTTGGGCTGTGCACTGTCAACAGGTGCAACACTGCTCATGCTC 1416
 DB 341 ATTGATCTCTGATCTTGGGCTGTGCACTGTCAACAGGTGCAACACTGCTCATGCTC 400

QY 1417 ACACTGCCCCCTGTGAGATCTTCCCTGGGCGCTCTGCCCCCTGCTTCCACAGACAC 1476

DB 401 AACTGCCCCCTGTGAGATCTTCCCTGGGCGCTCTGCCCCCTTCCACAGACAC 460

QY 1477 TTTCTTGGGCTTAAGGCTTCTCTCTCAGGACCTTAATTTGACCAACCAACCTGGGCT 1536
 DB 461 TTTCTTGGGCTTAAGGCTTCTCTCTCAGGACCTTAATTTGACCAACCAACCTGGGCT 520

QY 1537 TCAGCCATCATCATGTGGGCACTGGAGCTGGGGGTGCAATGGGGCTGTCTACCTTGGCCAC 1596
 DB 521 TCAGCCATCATCATGTGGGCACTGGAGCTGGGGGTGCAATGGGGCTGTCTACCTTGGCCAC 580

QY 1597 ACATCTCCAGCCAGGAGGCGCCCTGCCAGCTTCAATTACAGACTGACTCTCTCACC 1656
 DB 581 ACATCTCCAGCCAGGAGGCGCCCTGCCAGCTTCAATTACAGACTGACTCTCTCACC 640

QY 1657 TTCCCCCTGTGCTGCAGAGCTGAACATAGACTTGCATTTGATGATCTGAGAGTGTCA 1716
 DB 641 TTCCCCCTGTGCTGCAGAGCTGAACATAGACTTGCATTTGATGATCTGAGAGTGTCA 700

QY 1717 CATGGAGGTGTTATGGCA 1734
 DB 701 CATGGAGGTGTTATGGCA 718

RESULT 36
 AW245401/c 588 bp mRNA linear EST 07-JAN-2000
 LOCUS 2822785.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822785 3'
 DEFINITION mRNA sequence.

ACCESSION AW245401
 VERSION AW245401.1 GI:6588394
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 588)
 NIH-MGC http://mgi.mgi.nhl.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Other ESTs: 2822785.5prime
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCl/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Berkeley MGC sequencing project
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://www.lnl.gov/bprp/image/image.html
 Base Calling / Quality Scores: PHRED from University of Washington Genome Center. Vector Trimming: cross match from University of Washington Genome Center PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center: http://www.genome.washington.edu
 Polyadenylation: Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this cDNA insert was polyadenylated.
 Plate: LNCM10 row: F column: 2
 High quality sequence stop: 521.
 Location/Qualifiers
 1..588
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2822785"
 /tissue_type="small cell carcinoma"
 /cell_line="MGC3"
 /lab_host="DH10B (phage-resistant)"
 /clone_1lb="NIH_MGC_7"
 /note="Organ: lung; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(G). Size-selected >500bp for average

Insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

```
Query Match      29.5%; Score 555; DB 2; Length 588;
Best Local Similarity 100.0%; Pred. No. 2e-271;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1329 TACACACTACCAAGAGGCCCCCTGACATTTATCTTATCTTGGGCTGTCTGACATG 1388
DB 565 TACACACTACCAAGAGGCCCCCTGACATTTATCTTATCTTGGGCTGTCTGACATG 506
QY 1389 TCACAGGTGACACACTGCTGATGTCACATGCCCCCTGATGATCTTCCCTGGGCTT 1448
DB 505 TCACAGGTGACACACTGCTGATGTCACATGCCCCCTGATGATCTTCCCTGGGCTT 446
QY 1449 CTGCCCCGCTGCTGCTTCCAGACACACTTCTTGGGCTTAAAGGCTTCTCTCAGGACC 1508
DB 445 CTGCCCCGCTGCTGCTTCCAGACACACTTCTTGGGCTTAAAGGCTTCTCTCAGGACC 386
QY 1509 TCTAATTTAGCAACAACTGAGCTTACGCCATCATGAGGCACTGAGAGTGGGT 1568
DB 385 TCTAATTTAGCAACAACTGAGCTTACGCCATCATGAGGCACTGAGAGTGGGT 326
QY 1569 GCACATGGGGGCTGCTACCTTGGCCACATCTCCAGCCAGGCCCCCTGCCAGCT 1628
DB 325 GCACATGGGGGCTGCTACCTTGGCCACATCTCCAGCCAGGCCCCCTGCCAGCT 266
QY 1629 TCAATTTACAGACCTGACTCTCTCCTCACTTCCCCCTGCTGTCCAGAGCTGAACATGAG 1688
DB 265 TCAATTTACAGACCTGACTCTCTCCTCACTTCCCCCTGCTGTCCAGAGCTGAACATGAG 206
QY 1689 TTGCACCTGGATGTCACTTGAAGTGTCACTGAGAGTGTATGAGGACATCATCAAGG 1748
DB 205 TTGCACCTGGATGTCACTTGAAGTGTCACTGAGAGTGTATGAGGACATCATCAAGG 146
QY 1749 CCTACCTGTTGCAATGGGGCCAAACAGATTAACAGGCACTTCTTGGAAGGAATGCA 1808
DB 145 CCTACCTGTTGCAATGGGGCCAAACAGATTAACAGGCACTTCTTGGAAGGAATGCA 86
QY 1809 AAGGCTTTGGGGGTGATGAAAGACCTTTTACCAATGATACCAATTAACCTGCCCTGA 1868
DB 85 AAGGCTTTGGGGGTGATGAAAGACCTTTTACCAATGATACCAATTAACCTGCCCTGA 26
QY 1869 AAGGGCATAGTGGG 1883
DB 25 AAGGGCATAGTGGG 11

RESULT 37
CB990003 817 bp mRNA linear EST 01-MAY-2003
LOCUS AGENCOURT 13890677 NIH MGC 147 Homo sapiens cDNA clone
DEFINITION IMAGE:30347629 5', mRNA sequence.
ACCESSION CB990003
VERSION CB990003.1 GI:30284523
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 817)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgabs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: NDAM389 row: a column: 14
High quality sequence stop: 566.
Location/Qualifiers
1. 817

FEATURES

source

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30347629"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 147"
/notes="Organ: Placenta; Vector: pBluescriptR; Site_1:
all-XhoI; Site_2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to R05. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."
```

ORIGIN

```
Query Match      28.9%; Score 545; DB 6; Length 817;
Best Local Similarity 99.8%; Pred. No. 2.5e-266;
Matches 595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 TAATATAGAGAGTGGCCAGAAAGATCCAAACAGTGGCTGCGGCGGCCAGAG 65
DB 33 TAATATAGAGAGTGGCCAGAAAGATCCAAACAGTGGCTGCGGCGGCCAGAG 92
QY 66 TCATGCAAGCCAGAAATCGGCGGCTTCTGAGCTTGTTCGCTTCCCTCCCGGAAAT 125
DB 93 TCATGCAAGCCAGAAATCGGCGGCTTCTGAGCTTGTTCGCTTCCCTCCCGGAAAT 152
QY 126 GGGGCTATCCGGGTGCAACCCCGGCGGCTGCGGAGAGAGATGCTTGGACTATTA 185
DB 153 GGGGCTATCCGGGTGCAACCCCGGCGGCTGCGGAGAGAGATGCTTGGACTATTA 212
QY 186 CGGAGTCTGTGCTTCACTGATGTTGAGAGTGTGAGCGGCACTGACCGAGTGGCA 245
DB 213 CGGAGTCTGTGCTTCACTGATGTTGAGAGTGTGAGCGGCACTGACCGAGTGGCA 272
QY 246 GCTGAGACTTCTGAGCTTCTGCTGAGATGAGCTCTTGGCGCGGAGAGCTTACCCG 305
DB 273 GCTGAGACTTCTGAGCTTCTGCTGAGATGAGCTCTTGGCGCGGAGAGCTTACCCG 332
QY 306 GGGCCGAGCGGCTTGAAGCTCTGCTGAGATGAGCTCTTGGCGCGGAGAGCTTACCCG 365
DB 333 GGGCCGAGCGGCTTGAAGCTCTGCTGAGATGAGCTCTTGGCGCGGAGAGCTTACCCG 392
QY 366 CAACCTGCGGCTGCTGAGGCACTCTGCGCGGCTGAGCGGCGGCACTGCTGCGGCA 425
DB 393 CAACCTGCGGCTGCTGAGGCACTCTGCGCGGCTGAGCGGCGGCACTGCTGCGGCA 452
QY 426 CTTGCGCGCAAGCGGCGGCGGCACTGCTGCGAGAGCTATAGTATAGCACTTCAG 485
DB 453 CTTGCGCGCAAGCGGCGGCGGCACTGCTGCGAGAGCTATAGTATAGCACTTCAG 512
QY 486 CTTTCAAGAGAGAGAGAGTGTGCTGCGGCTGCGGCTGCGGCACTGCTGCGGCA 545
DB 513 CTTTCAAGAGAGAGAGAGTGTGCTGCGGCTGCGGCTGCGGCACTGCTGCGGCA 572
QY 546 TTCTCAGAGAGGCTAGTGGAGAGACAGAGCTTCCCGGCAACCAAGCGGCGAGCGGCA 601
DB 573 TTCTCAGAGAGGCTAGTGGAGAGACAGAGCTTCCCGGCAACCAAGCGGCGAGCGGCA 628

RESULT 38
CB994488
```

LOCUS CB994488 780 bp mRNA linear EST 01-MAY-2003
 DEFINITION IMAGE:13620146 NIH_MGC_148 Homo sapiens cDNA clone
 ACCESSION CB994488
 VERSION CB994488.1 GI:30289008
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 780).
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM551 row: P column: 13
 High quality sequence stop: 576.
 Location/Qualifiers
 1. 780
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3033396"
 /tissue_type="pre-eclamplic placenta"
 /lab_host="DH10B Tona"
 /clone_1ib="NIH MGC 148"
 /note="Organ: placenta; Vector: pBluescriptR; Site_1:
 all-XhoI; Site_2: BamI; library is oligo-dT primed and
 directionally cloned using primer
 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert
 size 2.3 kb and normalized to ROT 5. This is a primary
 library enriched for full-length clones and constructed
 using the Cap-trapper method (Carninci, in preparation).
 Library constructed by M. Brownstein (NIH/NHGRI,
 National Institutes of Health). Note: this is a NIH_MGC
 Library."

ORIGIN

Query Match 28.4%; Score 535; DB 6; Length 780;
 Best Local Similarity 99.7%; Pred. No. 3.2e-261;
 Matches 635; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

35 AAAACAAGTGGCTGCGGCGCTGCCAGAGTATCGAGCAGCAATCTGCGGGGTTTC 94
 28 AAAACAAGTGGCTGCGGCGCTGCCAGAGTATCGAGCAGCAATCTGCGGGGTTTC 87
 95 TGAAGTTGTTCCGCTCTCTCCCGGGAATGCGGCTATCGGAGTGAACCCGAGCCCGT 154
 88 TGAAGTTGTTCCGCTCTCTCCCGGGAATGCGGCTATCGGAGTGAACCCGAGCCCGT 147
 155 GCTGGAGAGAGAGTGTGCTGAGCTACTACGGAATGCTGTGCTTCAACCGATGTTG 214
 148 GCTGGAGAGAGAGTGTGCTGAGCTACTACGGAATGCTGTGCTTCAACCGATGTTG 207
 215 AGGTGAGTGGCGGAGCACTGACGAGTGTGAGTCTGAGCTTCTGCTGGATG 274
 208 AGGTGAGTGGCGGAGCACTGACGAGTGTGAGTCTGAGCTTCTGCTGGATG 267
 275 AGGCTCTGCGGCGCGGAGGCTTACCGGAGCGGAGCGGAGCTTACGAGCTCTGCTGG 334
 268 AGGCTCTGCGGCGCGGAGGCTTACCGGAGCGGAGCGGAGCTTACGAGCTCTGCTGG 327
 335 AGCTGAGAGCGCGGCGGAGTGGGAGAGCAACTGCGGCTGTGGGCACTCTCTGC 394

Db 328 AGCTGAGAGCGCGGCGGAGTGGGAGAGCAACTGCGGCTGTGGGCACTCTCTGC 387
 Qy 395 GCGTGTGTCGCCCGCAGCACTGTCGCGCACTTGGCGGAGCGCGCGGCAAGTGT 454
 Db 388 GCGTGTGTCGCCCGCAGCACTGTCGCGCACTTGGCGGAGCGCGCGGCAAGTGT 447
 Qy 455 CTCAGAGAGCTATGCTATGCTGACCTCTTCAAGAGAGCAAGAGTTACTGTGC 514
 Db 448 CTCAGAGAGCTATGCTATGCTGACCTCTTCAAGAGAGCAAGAGTTACTGTGC 507
 Qy 515 GTGCGCGTGGAGTCAAGCACTTTCGCAATTTTCAGAGGAGTCAAGAGAGAGCT 574
 Db 508 GTGCGCGTGGAGTCAAGCACTTTCGCAATTTTCAGAGGAGTCAAGAGAGAGCT 567
 Qy 575 CCCCCCAACCAAGCGAGCGCGGAGTGGGAGCGGCCAGTGTGTCAGACGCG 634
 Db 568 CCCCCCAACCAAGCGAGCGCGGAGTGGGAGCGGCCAGTGTGTCAGACGCG 627
 Qy 635 GCGGAGAGAGGCGCGGAGCGCGGAGTGGGAGCGGCCAGTGTGTCAGACGCG 671
 Db 628 GCGGAGAGAGGCGCGGAGCGCGGAGTGGGAGCGGCCAGTGTGTCAGACGCG 664

RESULT 39
 BE674778/c 533 bp mRNA linear EST 08-SEP-2000
 LOCUS BE674778
 DEFINITION 7e98906.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:3293242 3',
 mRNA sequence.
 ACCESSION BE674778
 VERSION BE674778.1 GI:10035246
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 533)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
 M.D., Louis M. Staudt, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL, send email to:
 info@image.llnl.gov
 Seq primer: -40up from Gibco
 High quality sequence stop: 460.
 Location/Qualifiers
 1. 533
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3293242"
 /tissue_type="B-cell, chronic Lymphocytic leukemia"
 /lab_host="DH10B"
 /clone_1ib="NCI_CGAP CLL1"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGAGTGGAGCGGCGGAGTGTGTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT7T3 vector.
 Library is normalized, and was constructed by Bento
 Soares and M. Fatima Bonaldo."

FEATURES

source

ORIGIN

Query Match

28.3%; Score 532; DB 2; Length 533;

REFERENCE 1 (bases 1 to 586)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. 586
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="STM08261"
/tissue_type="stomach mucosa"
/clone_lib="Sugano cDNA library, stomach mucosa"

ORIGIN
Query Match 28.1%; Score 530; DB 5; Length 586;
Best Local Similarity 99.8%; Pred. No. 1.1e-258;
Matches 580; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 40 AAGTGGCTGGCGCGCTGCGCCAGAGTCAATCGGACGCCAGAAATCTGCGCGGTTCTGACC 99
DB 1 AAGTGGCTGGCGCGCTGCGCCAGAGTCAATCGGACGCCAGAAATCTGCGCGGTTCTGACC 60

QY 100 TTGTTTCCGCTCCTCTCCCGGGAAATGGGCTTATCCGGGTCCACCCCGGCGCCCTGCTGG 159
DB 61 TTGTTTCCGCTCCTCTCCCGGGAAATGGGCTTATCCGGGTCCACCCCGGCGCCCTGCTGG 120

QY 160 GAGGAGATAGAGTGCCTGACCTACTAGGGATCTGCTGCTTACCGTATGTTGAGAGTG 219
DB 121 GAGGAGATAGAGTGCCTGACCTACTAGGGATCTGCTGCTTACCGTATGTTGAGAGTG 180

QY 220 GTGGGCGGGCACTGACCGAGTGCAGTGGAGCTCTTGCGCTTCTGCTGATGAGGCT 279
DB 181 GTGGGCGGGCACTGACCGAGTGCAGTGGAGCTCTTGCGCTTCTGCTGATGAGGCT 240

QY 280 CCGTGGCGCGCGGAGGCTTGAAGCCCGGCGCGGAGGCGGCTTGAAGCTCTGCTGAGAGTG 339
DB 241 CCGTGGCGCGCGGAGGCTTGAAGCCCGGCGCGGAGGCGGCTTGAAGCTCTGCTGAGAGTG 300

QY 340 GAGCGCGCGGAGTGCAGTGCAGAGCAACTGCTGCTGAGGAGCACTCTGCGCGTG 399
DB 301 GAGCGCGCGGAGTGCAGTGCAGAGCAACTGCTGCTGAGGAGCACTCTGCGCGTG 360

QY 400 CTGGCCCGCCAGCACTGCTGCGCACTTGGCGCGCAAGCGGCGCGGCGCACTGCTTCCA 459
DB 361 CTGGCCCGCCAGCACTGCTGCGCACTTGGCGCGCAAGCGGCGCGGCGCACTGCTTCCA 420

QY 460 GAACGCTATAGCTATGAGCACTTCAAGTCTTCAAAGAGACAAAGGATAGTCCCTGCG 519
DB 421 GAACGCTATAGCTATGAGCACTTCAAGTCTTCAAAGAGACAAAGGATAGTCCCTGCG 480

QY 520 CGTGGCAGTCAAGCACTTCTGCAAAATCTCAGCAGGAGTCAAGGAGAGAGAGGCTCCCC 579
DB 481 CGTGGCAGTCAAGCACTTCTGCAAAATCTCAGCAGGAGTCAAGGAGAGAGAGGCTCCCC 540

QY 580 CCAACCAAGCGGAGCGGCGGAGTCTGGGCGCGGCGGCGGAGTGG 620
DB 541 CCAACCAAGCGGAGCGGCGGAGTCTGGGCGCGGCGGCGGAGTGG 581

RESULT 42
BP253687 582 bp mRNA linear EST 15-SEP-2004
LOCUS BP253687 Sugano cDNA library, kidney epithelial cell Homo sapiens
DEFINITION CDNA clone HRC04630, mRNA sequence.
ACCESSION BP253687
VERSION BP253687.1 GI:52135968

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. 582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC04630"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"

ORIGIN
Query Match 27.7%; Score 522; DB 5; Length 582;
Best Local Similarity 99.8%; Pred. No. 1.4e-254;
Matches 572; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 GTAATATATAGAGAGGCTGCAGAAAGATCAAAACAAGTGGCTGGCGCGCTGCCAGAGA 64
DB 1 GTAATATATAGAGAGGCTGCAGAAAGATCAAAACAAGTGGCTGGCGCGCTGCCAGAGA 60

QY 65 GTATATCGAGCGCGAGATGCGCGGCTTCTAGAGTGTTCGCGCTCCCTCCCGGGGAA 124
DB 61 GTATATCGAGCGCGAGATGCGCGGCTTCTAGAGTGTTCGCGCTCCCTCCCGGGGAA 120

QY 125 TGGCGCTATCCGAGTTCGACCCCGGCGCGGCTGCTGGAGAGAGATGAGTCTGACTACT 184
DB 121 TGGCGCTATCCGAGTTCGACCCCGGCGCGGCTGCTGGAGAGAGATGAGTCTGACTACT 180

QY 185 ACGGAGTCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGCACTGACCGAGTGG 244
DB 181 ACGGAGTCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGCACTGACCGAGTGG 240

QY 245 AACTGAGCTCTGCGCTTCTGCTGATGAGGCTCTGGGCGCGCGGAGGCTTAGGCC 304
DB 241 AACTGAGCTCTGCGCTTCTGCTGATGAGGCTCTGGGCGCGCGGAGGCTTAGGCC 300

QY 305 GGGCCCGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGCAGCGAGA 364
DB 301 GGGCCCGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGGCGAGTGCAGCGAGA 360

QY 365 GCAACTGCGCTGTGCTGGGCAACTCTCTGCGCTGTGCGCGGCGCACTGCTGCGCG 424
DB 361 GCAACTGCGCTGTGCTGGGCAACTCTCTGCGCTGTGCGCGGCGCACTGCTGCGCG 420

QY 425 AACTGCGCGGCAAGCGCGCGCGCGGCAAGTCTTCCGAAGAGCTATAGCTATGCACTGCA 484
DB 421 AACTGCGCGGCAAGCGCGCGCGCGGCAAGTCTTCCGAAGAGCTATAGCTATGCACTGCA 480

QY 485 GCTCTTCAAGAGAGCAGAGGATAGCTGCGCTGCGCTGCGAGTCAACAGATTCTGCAA 544
DB 481 GCTCTTCAAGAGAGCAGAGGATAGCTGCGCTGCGCTGCGAGTCAACAGATTCTGCAA 540

QY 545 ATTCTCAGCAGGATCAGTGGAGACAGAGGCTCCC 577
DB 541 ATTCTCAGCAGGATCAGTGGAGACAGAGGCTCCC 573

RESULT 43

BP306442 569 bp mRNA linear EST 17-SEP-2004
LOCUS BP306442 Sugano cDNA library, macrophage Homo sapiens cDNA clone
DEFINITION MG07447, mRNA sequence.
ACCESSION BP306442
VERSION BP306442.1 GI:52235402
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 569)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
source
1. 569
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MG07447"
/cell_type="macrophage"
/clone_lib="Sugano cDNA library, macrophage"

ORIGIN
Query Match 27.7%; Score 521; DB 5; Length 569;
Best Local Similarity 100.0%; Pred. No. 4.5e-254; Indels 0; Gaps 0;
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 451 GTGTCTTCAGAACGCTATAGCTATGCGACCTCCAGCTCTTCAAGAGAGAGAGGTAGC 510
DB 49 GTGTCTTCAGAACGCTATAGCTATGCGACCTCCAGCTCTTCAAGAGAGAGAGGTAGC 108
QY 511 TGGCGTTCGCGTGGCGAGTCAAGAGTTCGCAATTCTCAGCAGGGTCACTGGAGACA 570
DB 109 TGGCGTTCGCGTGGCGAGTCAAGAGTTCGCAATTCTCAGCAGGGTCACTGGAGACA 168
QY 571 GGCTCCCGCCCAACGAGCGGAGCGGAGTCCGGGCGCGCCAGTGTGGTCCAGA 630
DB 169 GGCTCCCGCCCAACGAGCGGAGCGGAGTCCGGGCGCGCCAGTGTGGTCCAGA 228
QY 631 CGGCGGAGAGAGGGGCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 690
DB 229 CGGCGGAGAGAGGGGCCCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 288
QY 691 TCTGAAGGCAAGTGAATCTGTGATCCGCTCCGGTTCGAGCAGAGTACTGGAGCAT 750
DB 289 TCTGAAGGCAAGTGAATCTGTGATCCGCTCCGGTTCGAGCAGAGTACTGGAGCAT 348
QY 751 GGGCGAGCTTTGAGAGAGGGGCGTGGCATCCGGGCGCGCCGAGGCGCTGGGCGGAGCATG 810
DB 349 GGGCGAGCTTTGAGAGAGGGGCGTGGCATCCGGGCGCGCCGAGGCGCTGGGCGGAGCATG 408
QY 811 GAGGTGTTTGGGAGAGGCGCAGCGAGTGGTGGCTCAAGAGGAGCTGGGCTGTGGTTGT 870
DB 409 GAGGTGTTTGGGAGAGGCGCAGCGAGTGGTGGCTCAAGAGGAGCTGGGCTGTGGTTGT 468
QY 871 GACATCAAGTTCTCAGAGCTCTCTATCTGGAAGCTTCTGGGGCGACTACCTGAGTGGC 930
DB 469 GACATCAAGTTCTCAGAGCTCTCTATCTGGAAGCTTCTGGGGCGACTACCTGAGTGGC 528
QY 931 GCCCTGCTGAGAGCGCTGGGGGCGTTCCTGACTGAGAGC 971
DB 529 GCCCTGCTGAGAGCGCTGGGGGCGTTCCTGACTGAGAGC 569

RESULT 44
LOCUS B1827768 832 bp mRNA linear EST 04-OCT-2001
DEFINITION 603075430F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5167254 5',
mRNA sequence.
ACCESSION B1827768
VERSION B1827768.1 GI:15939318
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 832)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L14M1415 row: 1 column: 07
High quality sequence stop: 670.
Location/Qualifiers
source
1. 832
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5167254"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: Brain; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: BcoRI (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (BcoRI site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

ORIGIN
Query Match 27.5%; Score 517; DB 4; Length 832;
Best Local Similarity 99.8%; Pred. No. 5.1e-252;
Matches 567; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 60 CAGAGTCAATCGAGCGCAGAAATCTGGCGGCTTCTGAGGTTTCCCTCCCTCC 119
DB 1 CAGAGTCAATCGAGCGCAGAAATCTGGCGGCTTCTGAGGTTTCCCTCCCTCC 60
QY 120 GGGAGTGGGCTATCCGGGGTCAAGCCCGGCGCGCTGGGAGAGAGTGAAGCTCGA 179
DB 61 GGGAGTGGGCTATCCGGGGTCAAGCCCGGCGCGCTGGGAGAGAGTGAAGCTCGA 120
QY 180 CTACTAGGAGATCTGCTGCTTCAACCGTATGTTGAGAGTGGTGGGCGGCACTGACGA 239
DB 121 CTACTAGGAGATCTGCTGCTTCAACCGTATGTTGAGAGTGGTGGGCGGCACTGACGA 180
QY 240 GTGCGAGCTGGAAGCTCTGGCTTTTCTGTGATGAGGCTCTTGGGCGCGGAGGCTT 239
DB 181 GTGCGAGCTGGAAGCTCTGGCTTTTCTGTGATGAGGCTCTTGGGCGCGGAGGCTT 240
QY 300 AGCCGAGGCGCGGAGGCGCTAGAGCTCTGCTGAGAGCTGAGAGCGCGGAGGAGTGGG 359
DB 241 AGCCGAGGCGCGGAGGCGCTAGAGCTCTGCTGAGAGCTGAGAGCGCGGAGGAGTGGG 300
QY 360 CGAGAGCAACTCTGGGCTCTGGGGCACTCTGCGCGTGTGGCCCGCAGCACTGTCT 419

Db	301	CGAAGCAACCTGGGGGCTGCTGGGGCAACCTCTGGAGGTGCTGGGCGGCCCAACCTGCT	360
Qy	420	GCCGCACTGGCGCGCAAGCGGCGCGCGCCAGTGTCTTCAGAAAGCTATAGCTATGCGAC	479
Db	361	GCCCCACCTGGCGCGCAAGCGGCGCGCGCCAGTGTCTTCAGAAAGCTATAGCTATGCGAC	420
Qy	480	CTCAGAGCTTTCAAAGAGGAGACAGAGGGTGTGCGCGTGGCGCGGAGCTGAAGAGTTTC	539
Db	421	CTCAGAGCTTTCAAAGAGGAGACAGAGGGTGTGCGCGTGGCGCGGAGCTGAAGAGTTTC	480
Qy	540	TGCAAAATTCTCAGAGAGGTTCAGTGGGAGACAGAGCTCCCCCAACCAAGCGGACGCGCG	599
Db	481	TGCAAAATTCTCAGAGAGGTTCAGTGGGAGACAGAGCTCCCCCAACCAAGCGGACGCGCG	540
Qy	600	GAGTCGGGGCGCGGCCCAAGTGGTGTGCC	627
Db	541	GAGTCGGGGCGCGGCCCAAGTGGTGTGCC	568
RESULT 45			
BP211545			
LOCUS			
DEFINITION			
BP211545 Sugano cDNA library, cerebrum Homo sapiens cDNA clone			
ACCESSION			
BP211545			
VERSION			
BP211545.1 GI:52084436			
KEYWORDS			
EST.			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE			
AUTHORS			
Suzuki,Y., Yamaeishi,R., Shiota,M., Sakakibara,Y., Chiba,J.,			
1 (bases 1 to 580)			
Mizushima-Sugano,J., Nakai,K. and Sugano,S.			
Sequence comparison of human and mouse genes reveals a homologous			
block structure in the promoter regions			
Genome Res. 14 (9), 1711-1718 (2004)			
CONTACT: Yutaka Suzuki			
Department of Virology			
Institute of Medical Science, University of Tokyo			
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan			
Email: yusuzuki@ims.u-tokyo.ac.jp.			
JOURNAL			
COMMENT			
Location/Qualifiers			
1..580			
FEATURES			
Source			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CBR01216"			
/issue_type="cerebrum"			
/clone_11b="Sugano cDNA library, cerebrum"			
ORIGIN			
Query Match			
Best Local Similarity 99.8%; Pred. No.5.2e-251; Length 580;			
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	41	AGTGGCTGCGGCGCGCTGCGCCAGAGTCAATCGGACCGCAAAATTCGCGGTTCTGACT	100
Db	15	AGTGGCTGCGGCGCGCTGCGCCAGAGTCAATCGGACCGCAAAATTCGCGGTTCTGACT	74
Qy	101	TGTTCCGCTCTCCCTCCCGGGGAATGGCGCTATCCGGGTGCAACCCGCGCCGCTGCTGGG	160
Db	75	TGTTCCGCTCTCCCTCCCGGGGAATGGCGCTATCCGGGTGCAACCCGCGCCGCTGCTGGG	134
Qy	161	AGGAGGATGATGCTGCTGAGACTATCGGGAATGCTGCTTCACCGTATGTTTCGAGGTGG	220
Db	135	AGGAGGATGAGTGGCTGAGACTATCGGGAATGCTGCTTCACCGTATGTTTCGAGGTGG	194
Qy	221	TGGGCGGGCAACTGACCGAGTGGCGAGTCTGAGACTCTGGACCTTTTCTGCTGATGAGCTTC	280
Db	195	TGGGCGGGCAACTGACCGAGTGGCGAGTCTGAGACTCTGGACCTTTTCTGCTGATGAGCTTC	254
Qy	281	CTGGGCGCGCGGAGGCTTAAAGCTCCCGGGCGCGACGCGCTTAAGAGCTCTGCTGAGGCTGG	340

Db	255	CTGGCGCGCGCGAGGCTTAGCTAGCCCGGCGCGAGCGGCTTAGAGTCTCTGCTGAGAGCTGG	314
QY	341	AGCGCCGCGGCGCACTGTCGGCGAGCACTTCGCGCTGTGGGGCAACTCTCTGCGCCGTGC	400
Db	315	AGCCCGCGGCGAGTGGCGACAGAGCAACTTCGCGCTGTGGGGCAACTCTCTGCGCGTGC	374
QY	401	TGGCGCGCGACGACTGTGCGGACCTGGCGCGCAACGCGCGCGCGCGAGTGTCTCCAG	460
Db	375	TGGCGCGCGACGACTGTGCGGACCTGGCGCGCGCAACGCGCGCGCGAGTGTCTCCAG	434
QY	461	AACGCTATAGCTATGCGACTCTCAGCTTTCTAAAGAGGACAGAGGGTAGTCTCGTGCC	520
Db	435	AAGCTATAGCTATGCGACTCTCAGCTTTCTAAAGAGGACAGAGGGTAGTCTCGTGCC	494
QY	521	GTGGCGAGTCAAGCAATTCTGCAAAATTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCC	580
Db	495	GTGGCGAGTCAAGCAATTCTGCAAAATTCTCAGCAGGGTCACTGGGAGACAGGCTCCCCC	554
QY	581	CAACCAAGCGGCGAGCGCGGAGTGGG	606
Db	555	CAACCAAGCGGCGAGCGCGGAGTGGG	580

```

RESULT 46
BE883568
LOCUS
DEFINITION
601508040F1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3909753 5',
ACCESSION
BE883568
VERSION
BE883568.1 GI:10332344
KEYWORDS
EST.
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 793)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
AUTHORS
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
COMMENT
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: LLM9723 row: p column: 10
High quality sequence, stop: 678.
Location/Qualifiers
1..793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3909753"
/rissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb. "
ORIGIN
Query Match 27.3%; Score 515; DB 2: Length 793;
Best Local Similarity 100.0%; Prid. No. 5.3e-251; Indels 0; Gaps 0
Matches 515; Conservative 0; Mismatches 0;
485 GCTCTTCAAGAGACAGAGGGTGTGCTGCGCTCGGCACTCAAGCAGTTCTGCAG 544
|||||
126 GCTTTCAAGAGACAGAGGGTGTGCTGCGCTCGGCACTCAAGCAGTTCTGCAG 185

```

```
QY 545 ATTCTCAGCAGGGTCACTGAGAGACAGAGCTCCCCCAACCAAGCGGAGGAGAGTC 604
DB 186 ATTCTCAGCAGGGTCACTGAGAGACAGAGCTCCCCCAACCAAGCGGAGGAGAGTC 245
QY 605 GGGGCGGCGCCCAAGTGTGTGTGTCCAGAGCGGGCGGAGAGGGGCCCCAGCCGACCCAGC 664
DB 246 GGGGCGGCGCCCAAGTGTGTGTGTCCAGAGCGGGCGGAGAGGGGCCCCAGCCGACCCAGC 305
QY 665 AGCAGTCAGAGCCCGCAGACCTTCTCTGAAAGGCAAAAGTACCTGTGACATCCGGCTCC 724
DB 306 AGCAGTCAGAGCCCGCAGACCTTCTCTGAAAGGCAAAAGTACCTGTGACATCCGGCTCC 365
QY 725 GGGTTCAGAGAGACTACTGCGAGCAGTGGCCAGCCTTGAGCAGAGCGGCTGGCATCCGAGC 784
DB 366 GGGTTCAGAGAGACTACTGCGAGCAGTGGCCAGCCTTGAGCAGAGCGGCTGGCATCCGAGC 425
QY 785 GGGCCCGAGGCGCTGGCGCGGCACTGAGAGTGTGTGGGCGAGGCCACCGCAGTGTGGCT 844
DB 426 GGGCCCGAGGCGCTGGCGCGGCACTGAGAGTGTGTGGGCGAGGCCACCGCAGTGTGGCT 485
QY 845 CAAGGAGCTGGGCTGTGGTGTGTGACATCAAGTCTGACAGCTCTCTATCTGGAGC 904
DB 486 CAAGGAGCTGGGCTGTGGTGTGTGACATCAAGTCTGACAGCTCTCTATCTGGAGC 545
QY 905 CTTTCTGGGCGACTACTGAGTGGCGCCCTGCTGCAAGCCCTGCGGGCGTGTCTCTGA 964
DB 546 CTTTCTGGGCGACTACTGAGTGGCGCCCTGCTGCAAGCCCTGCGGGCGTGTCTCTGA 605
QY 965 CTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCT 999
DB 606 CTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCT 640
```

RESULT 47

```
CD365564/c 651 bp mRNA linear EST 05-AUG-2004
LOCUS UT-H-FT2-bj-k-e-13-0-UI.s1 NCI CGAP FT2 Homo sapiens cDNA clone
DEFINITION UT-H-FT2-bj-k-e-13-0-UI 3', mRNA sequence.
ACCESSION CD365564
VERSION CD365564.1 GI:31149654
KEYWORDS EST.
```

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens [Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/cgap.html>
Seq primer: M13 FORWARD
POLYA=yes.

FEATURES

source

```
1. 651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-H-FT2-bj-k-e-13-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FT2"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: Scov I; Site_2: Not I;
NCI_CGAP_FT2 is a subtracted cDNA library constructed from
```

ORIGIN

Query Match 27.2%; Score 513; DB 6; Length 651;

Best Local Similarity 99.7%; Pred. No. 5.5e-250; Mismatches 0; Gaps 0;

Matches 613; Conservative 0; Indels 0; Gaps 0;

a pool of 81 RNA samples from Alveolar Macrophages challenged with different treatments. The RNA samples were a mixture of these conditions (lines refer to incubations following isolation by bronchoalveolar lavage) (some normal donor macrophages were cultured in some of the conditions, other donor macrophages in different conditions). The RNA samples were pooled for library construction. Control 0 hours; control 3 hours; control 24 hours; LPS 100 ng/mL, 3 hours; LPS 100 ng/mL, 24 hours; PMA 10 ng/mL, 3 hours; PMA 10 ng/mL, 24 hours; Staph aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Staph aureus moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 3 hours; Adenoviral vector (Ad5 CMV eGFP), moi 500, 24 hours; wt adenovirus moi 500, 3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS 3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3 hours; wt adenovirus + LPS 24 hours. The library was subcloned according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.

TAG TISSUE=Human Lung Alveolar Macrophage

TAG_LIB=UI-H-FT2

TAG_SEQ=GGCCATGCGC"

```
QY 1269 TGGCCCTTACAGCTCCGAGAGCGGGGCTGGGCTGTATCTCAGAAAGGAGGACAGC 1328
DB 633 TGGCCCTTACAGCTCCGAGAGCGGGGCTGGGCTGTATCTCAGAAAGGAGGACAGC 574
QY 1329 TACACACTACCAAGGCCCCCTTGCACATTGTATCTGTATCTTGGGCTGTCTGCACTG 1388
DB 573 TACACACTACCAAGGCCCCCTTGCACATTGTATCTGTATCTTGGGCTGTCTGCACTG 514
QY 1389 TCAAGGTGACACACTCGCTCATGTGTCACAGCCCTGTGATCTTCCCTGGGCT 1448
DB 513 TCAAGGTGACACACTCGCTCATGTGTCACAGCCCTGTGATCTTCCCTGGGCT 454
QY 1449 CTGCCCCCTGCTCTCCAGACACACTTCTTGGCTTAAAGGCTTCTCTCAGAGCC 1508
DB 453 CTGCCCCCTGCTCTCCAGACACACTTCTTGGCTTAAAGGCTTCTCTCAGAGCC 394
QY 1509 TCTAATTGACCAACCAACCACTGGGCTTCAAGCATCATGAGGCTGGGGT 1568
DB 393 TCTAATTGACCAACCAACCACTGGGCTTCAAGCATCATGAGGCTGGGGT 334
QY 1569 GCAATGGGCGCTGCACCTTGCACACATCTCAGACGACGAGGCGCTTCCAGCT 1628
DB 333 GCAATGGGCGCTGCACCTTGCACACATCTCAGACGAGGCGCTTCCAGCT 274
QY 1629 TCAATTACAGACTACTCTCTCACTTCCCTGCTGTCCAGAGCTGAACATAGAC 1688
DB 273 TCAATTACAGACTACTCTCTCACTTCCCTGCTGTCCAGAGCTGAACATAGAC 214
QY 1689 TTGCACTTGGATGTCACTGTGAGTGTCAATGGAGGTGTTAGGACGATCATCAAGG 1748
DB 213 TTGCACTTGGATGTCACTGTGAGTGTCAATGGAGGTGTTAGGACGATCATCAAGG 154
QY 1749 CCTACTGTGCAATGGGCGCAAAACAGTAACACGACCTTCTGGAAGGAATGCA 1808
DB 153 CCTACTGTGCAATGGGCGCAAAACAGTAACACGACCTTCTGGAAGGAATGCA 94
QY 1809 AAGCTTTGGGGGTGATGTAAGAAACCTTTTACAAATGATACCAATTAACCTCCCTGGA 1868
DB 93 AAGCTTTGGGGGTGATGTAAGAAACCTTTTACAAATGATACCAATTAACCTCCCTGGA 34
QY 1869 AAGGCAATGAGTGGG 1883
DB 33 AAGGCAATGAGTGGG 19
```



```
RESULT 48
BO923358          931 bp  mRNA  linear  EST 20-AUG-2002
LOCUS             AGENCOURT_8784534 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6376221
DEFINITION        5', mRNA sequence.
ACCESSION         BO923358
VERSION           BO923358.1 GI:22338389
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE             Unpublished (1999)
JOURNAL           Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Rubin Laboratory
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L1CM2558 row: a column: 22
                  High quality sequence stop: 511.
FEATURES
source            Location/Qualifiers
                  1..931
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6376221"
                  /tissue_type="normal pigmented retinal epithelium"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_43"
                  /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                  EcoRI; CDNA made by oligo-dt priming. Directionally
                  cloned into EcoRI/XhoI sites using the following 5'
                  adaptor: GGCACGAG(G). Library constructed by Ling Hong
                  in the laboratory of Gerald M. Rubin (University of
                  California, Berkeley) using ZAP-cDNA synthesis kit
                  (Stratagene) and Superscript II RT (Life Technologies).
                  Note: this is a NIH_MGC Library. |"
ORIGIN
Query Match      27.0%; Score 509; DB 5; Length 931;
Best Local Similarity 99.8%; Pred. No. 6.2e-248;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 115 CCCCCGGGAATGCGCTATCCGGGTGACACCCGGCCGCTGCGAGAGAGATGATGC 174
DB 1 CCCCCGGGAATGCGCTATCCGGGTGACACCCGGCCGCTGCGAGAGAGATGATGC 60
QY 175 CTGACTACTACGCGAGTGTCTGCTTACCGTATGTTGAGGTGTGGCGGCAACTG 234
DB 61 CTGACTACTACGCGAGTGTCTGCTTACCGTATGTTGAGGTGTGGCGGCAACTG 120
QY 235 ACCGAGTGCAGCTGAGCTCTGCGCTTTCTGCTGATAGAGCTCTTGCGCCGCCGGA 234
DB 121 ACCGAGTGCAGCTGAGCTCTGCGCTTTCTGCTGATAGAGCTCTTGCGCCGCCGGA 180
QY 295 GGGCTTAGCCCGGGCCGCGACGGGCTTAGAGCTCTGTGTGAGGTGAGCGCGGGGAG 354
DB 181 GGGCTTAGCCCGGGCCGCGACGGGCTTAGAGCTCTGTGTGAGGTGAGCGCGGGGAG 240
QY 355 TGGCGGAGAGCAACTGTCGCGCTGTCGGGCAACTCTCTGCGCGTGTGCGCCGACGAC 414
DB 241 TGGGAGAGAGCAACTGTCGCGCTGTCGGGCAACTCTCTGCGCGTGTGCGCCGACGAC 300
QY 415 CTGCTGCCGCACTGGCGCGCAAGCGGCGCGGCACTGTCTCCAGAACGCTATAGCTAT 474
```

```
DB 301 CTGCTGCCGCACTGGCGCGCAAGCGGCGCGGCACTGTCTCCAGAACGCTATAGCTAT 360
QY 475 GGCACCTCCAGCTCTTCAAAAGAGACAGAGGGTAGTGCCTGCGCGTCCGACATCAAC 534
DB 361 GGCACCTCCAGCTCTTCAAAAGAGACAGAGGGTAGTGCCTGCGCGTCCGACATCAAC 420
QY 535 AGTTCTGCAAAATTTCTCAGCAGGGTCACTGAGAGACAGCTCCCCCAACCAAGCGGCA 594
DB 421 AGTTCTGCAAAATTTCTCAGCAGGGTCACTGAGAGACAGGCTCCCCCAACCAAGCGGCA 480
QY 595 CGCGGAGTCCGGGCGCGCCAGTGTGTGTCAGACGGCGGAGAGAGGAGCCCGGAGCC 654
DB 481 CGCGGAGTCCGGGCGCGCCAGTGTGTGTCAGACGGCGGAGAGAGGAGCCCGGAGCC 540
QY 655 GCACCCGACAGCAGCTCAGA 674
DB 541 GCACCCGACAGCAGCTCAGA 560
```

```
RESULT 49
BE391120          654 bp  mRNA  linear  EST 21-JUL-2000
LOCUS             BE391120
DEFINITION        mRNA sequence.
ACCESSION         BE391120
VERSION           BE391120.1 GI:9336485
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE         NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS           National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE             Unpublished (1999)
JOURNAL           Contact: Robert Strausberg, Ph.D.
                  Email: cgabbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Ling Hong/Rubin Laboratory
                  DNA Sequencing by: InCyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L1CM259 row: 1 column: 20
                  High quality sequence stop: 617.
FEATURES
source            Location/Qualifiers
                  1..654
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:3607843"
                  /tissue_type="endometrium, adenocarcinoma cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /clone_lib="NIH_MGC_44"
                  /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
                  EcoRI; CDNA made by oligo-dt priming. Directionally
                  cloned into EcoRI/XhoI sites using the following 5'
                  adaptor: GGCACGAG(G). Library constructed by Ling Hong
                  in the laboratory of Gerald M. Rubin (University of
                  California, Berkeley) using ZAP-cDNA synthesis kit
                  (Stratagene) and Superscript II RT (Life Technologies)."
```

```
ORIGIN
Query Match      27.0%; Score 508; DB 2; Length 654;
Best Local Similarity 100.0%; Pred. No. 2e-247;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 451 GTGTCTCCAGAACGCTATAGTATGCACTCAGCTCTTCAAAAGAGACAGAGGGTAGC 510
DB 61 GTGTCTCCAGAACGCTATAGTATGCACTCAGCTCTTCAAAAGAGACAGAGGGTAGC 120
```



```

QY 511 TGGCGTGGCGGTGGAGTCAGCAGTTCGCAAAATTCAGCAGAGGTGAGTGGAGACA 570
Db 121 TGGCGTGGCGGTGGAGTCAGCAGTTCGCAAAATTCAGCAGAGGTGAGTGGAGACA 180
QY 571 GGGTCCCCCAACCAAGCGGCGAGCGGAGTTCGGGGCCGAGTGTGTGCCAGA 630
Db 181 GGGTCCCCCAACCAAGCGGCGAGCGGAGTTCGGGGCCGAGTGTGTGTGCCAGA 240
QY 631 CGGCGGCGAGAGGGGGCCCGACGCGCAGCAGAGTCAGAGCCCGCGAGCCTTCC 690
Db 241 CGGCGGCGAGAGGGGGCCCGACGCGCAGCAGAGTCAGAGCCCGCGAGCCTTCC 300
QY 691 TCTGAAGCAAGTGAAGTGAATCCGCGCTCGGGTTTGAGCAGAGTACTCGAGCAT 750
Db 301 TCTGAAGCAAGTGAAGTGAATCCGCGCTCGGGTTTGAGCAGAGTACTCGAGCAT 360
QY 751 GGGCCAGCCTTGAAGCAGAGGGCGTGGCATCCCGCGGCGCCAGCGCTGGCGCGAGCTG 810
Db 361 GGGCCAGCCTTGAAGCAGAGGGCGTGGCATCCCGCGGCGCCAGCGCTGGCGCGAGCTG 420
QY 811 GACGTGTTGGAGAGGCGACCGAGTGTGCGCTCAAGGGGACCTGGGCTGTGTGTGT 870
Db 421 GACGTGTTGGAGAGGCGACCGAGTGTGCGCTCAAGGGGACCTGGGCTGTGTGTGT 480
QY 871 GACATCAAGTTCAGAGCTCTCTATCTGAGCGCTTCTGGGCGCAGTACTGAGTGGC 930
Db 481 GACATCAAGTTCAGAGCTCTCTATCTGAGCGCTTCTGGGCGCAGTACTGAGTGGC 540
QY 931 GCCCTGCTGAGGCGCTTGGCGGGCGTGT 958
Db 541 GCCCTGCTGAGGCGCTTGGCGGGCGTGT 568

```

```

RESULT 50      506 bp      mRNA      linear      EST 21-DEC-2000
BF665646      602119860F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:427031 5',
LOCUS          mRNA sequence.
BF665646
ACCESSION      BF665646.1 GI:11939541
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS        National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE          Unpublished (1999)
JOURNAL
COMMENT        Contact: Robert Strauberg, Ph.D.
                Email: cgabbs-remail.nih.gov
                Tissue Procurement: ATCC
                cDNA Library Preparation: CLONTECH Laboratories, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
                DNA Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LNL at:
                http://image.llnl.gov
                Plate: LCM1098 row: g column: 16
                High quality sequence stop: 506.
                Location/Qualifiers

```

```

FEATURES
source
1..506
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:427031"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (TI phage-resistant)"
/clone_1ib="NIH_MGC_56"
/notes="Organ: Brain; Vector: pDNR-LIB (Clontech); Site_1:
S11 (ggcgccctggcg); Site_2: S11 (ggcattatggcg);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'

```

adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGGCGCGAGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```

ORIGIN
Query Match      26.9%; Score 506; DB 2; Length 506;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 424 CACCTGGCGCGAAGGGCGCGCGGCGAGTGTCTCCAGAAAGCTATAGTATGACACTCC 483
Db 1 CACCTGGCGCGAAGGGCGCGCGGCGAGTGTCTCCAGAAAGCTATAGTATGACACTCC 60
QY 484 AGCTCTTCAAGAAGAGCAGAGGGTAGCTCCGTCGCGCGGTGCGCAGTCAAGCAGTTCTGCA 543
Db 61 AGCTCTTCAAGAAGAGCAGAGGGTAGCTCCGTCGCGCGGTGCGCAGTCAAGCAGTTCTGCA 120
QY 544 AATTCTCAGCAGGGGTAGTGGAGAGAGAGCTCCCCCAACCAAGCGGCGAGT 603
Db 121 AATTCTCAGCAGGGGTAGTGGAGAGAGAGCTCCCCCAACCAAGCGGCGAGT 180
QY 604 CGGCGCGCGCGCGAGTGTGTGCGCAGAGCGCGCGGAGAGGGGGCCCGACCGGACCCGAG 663
Db 181 CGGCGCGCGCGCGAGTGTGTGCGCAGAGCGCGCGGAGAGGGGGCCCGACCGGACCCGAG 240
QY 664 CAGCAGTCAAGAGCCCGCAGAGCTTCTCTGAAAGGCAAGTGAACCTGTGACATCCGGCTC 723
Db 241 CAGCAGTCAAGAGCCCGCAGAGCTTCTCTGAAAGGCAAGTGAACCTGTGACATCCGGCTC 300
QY 724 CGGGTTCAGACAGTACTGCGAGCATGTGGGCCAGCTTGGAGCAGAGGCGTGGATCCCGG 783
Db 301 CGGGTTCAGACAGTACTGCGAGCATGTGGGCCAGCTTGGAGCAGAGGCGTGGATCCCGG 360
QY 784 CGGCGCGCGCGCGTGGCGCGGCGAGCTGAGCGTGTGGGCGAGGCGCAGCGAGTGGTGGC 843
Db 361 CGGCGCGCGCGCGTGGCGCGGCGAGCTGAGCGTGTGGGCGAGGCGCAGCGAGTGGTGGC 420
QY 844 TCAAGGAGCTTGGGCTGTGTGTGTGACATCAAGTTCAGAGTCTCTATCTGAGAC 903
Db 421 TCAAGGAGCTTGGGCTGTGTGTGTGACATCAAGTTCAGAGTCTCTATCTGAGAC 480
QY 904 GCCTTTCTGGGGGAGTACTGAGTGG 929
Db 481 GCCTTTCTGGGGGAGTACTGAGTGG 506

```

Search completed: March 24, 2005, 03:06:32
Job time : 6756.99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 23, 2005, 16:09:50 ; Search time 1091.22 seconds
(without alignments)
10215.030 Million cell updates/sec

Title: US-10-030-271-3
Perfect score: 1883
Sequence: 1 agcgatcaatcagagaag.....ctggaaggcataggctg99 1883

Scoring table: OLIGO_NTC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1883	100.0	1883	5	AAf27408 Human apo
2	1130	60.0	1230	6	ABA94362 Human APR
3	1130	60.0	1230	12	AD086891 Human tum
4	1105	58.7	2045	3	AAH33283 Human col
5	1105	58.7	2045	4	AAH33283 Human col
6	1078	57.2	1924	6	AAH40080 Human DED
7	1078	57.2	1924	6	AAH40080 Human DED
8	984	52.3	2044	6	AAH40080 Human DED
9	930	49.4	981	10	AD079259 Human DED
10	909	48.3	909	5	AAf27407 Human apo
11	707	37.5	847	11	AD130856 Human apo
12	707	37.5	1966	6	AAH62603 cDNA sequ
13	680	36.1	1570	4	AAH99646 Human pro
14	656	34.8	857	6	ABZ70266 ATP synth
15	421	22.4	626	4	AAH07925 Human gen
16	368	19.5	1042	4	AAI60747 Human pol
17	368	19.5	1067	5	AAI58961 Human pol
18	368	19.5	1067	5	ADQ99183 DNA encod
19	368	19.5	1067	5	ADQ99183 DNA encod
20	316	16.8	111084	12	ADQ18808 Human sof

21	300	15.9	451	9	ACH41209 Human foe
22	296	15.7	360	2	AAf25404 Human gen
23	268	14.2	603	5	AAf93937 Primer sp
24	252	13.4	303	6	AAH40075 Human DED
25	252	13.4	303	6	AAH40075 Human DED
26	237	12.6	441	6	ABH80091 Human ova
27	187	9.9	1084	6	ABH80091 Human ova
28	187	9.9	1084	6	ABH80091 Human ova
29	165	8.8	381	4	AAH12893 Human NS
30	155	5.1	484	9	ACH25056 Human adu
31	45	2.4	45	10	AD079268 Human DED
32	41	2.2	41	6	ABZ70271 ATP synth
33	37	2.0	1067	5	AAI58961 Human pol
34	37	2.0	1067	5	ADQ99183 DNA encod
35	37	2.0	1067	5	ADQ99183 DNA encod
36	29	1.5	594	13	AD055421 Novel hum
37	26	1.4	26	10	AD079267 Human DED
38	25	1.3	25	10	AD079264 Human DED
39	25	1.3	25	10	AD079263 Human DED
40	25	1.3	33	6	ABZ70269 ATP synth
41	24	1.3	24	6	ABZ70267 ATP synth
42	23	1.2	33	6	ABZ70270 ATP synth
43	23	1.2	23	10	AD079265 Human DED
44	22	1.2	23	10	AD079266 Human DED
45	22	1.2	838	5	ADL44878 Human ova
46	22	1.2	1339	6	ABH8384 Pain regu
47	22	1.2	2718	12	AD063628 Novel hum
48	21	1.1	21	5	AAf27413 Human apo
49	21	1.1	21	6	AAH40093 Human DED
50	21	1.1	21	6	AAH59075 Primer #2
51	21	1.1	23	5	AAf27416 Human ova
52	21	1.1	323	5	ADL39650 Human ova
53	21	1.1	2703	3	AAZ50617 Human cal
54	21	1.1	2922	3	AAZ50617 Human cal
55	21	1.1	3003	3	AAZ50615 Human cal
56	21	1.1	3234	3	AAZ51398 Human cal
57	21	1.1	3234	3	AAZ51398 Human cal
58	21	1.1	3234	3	AAZ51398 Human cal
59	21	1.1	3234	3	AAZ51398 Human cal
60	21	1.1	24902	12	AD156630 Human pol
61	20	1.1	20	5	AAf27414 Human apo
62	20	1.1	41	6	ABZ70272 ATP synth
63	20	1.1	121	12	ADK93823 Polynucle
64	20	1.1	267	3	AAf28175 Human sec
65	20	1.1	540	12	ACH78621 Human sec
66	20	1.1	1066	10	AD050943 Human gen
67	20	1.1	1512	6	ABK50658 Human HNI
68	20	1.1	2610	9	ACC59397 Microbial
69	20	1.1	2673	11	ABD04389 Pseudomon
70	20	1.1	2715	11	ABD04082 Pseudomon
71	20	1.1	2814	11	ABD04271 Pseudomon
72	20	1.1	31812	11	ACN44700 Human gen
73	20	1.1	115853	12	ADQ19685 Human gen
74	20	1.1	349989	10	AD086916 Human GPC
75	19	1.0	60	6	ABN86337 Human sp1
76	19	1.0	327	6	ABN86337 Human sp1
77	19	1.0	360	10	AD086916 Human sp1
78	19	1.0	360	10	AD086916 Human sp1
79	19	1.0	360	10	AD086916 Human sp1
80	19	1.0	360	10	AD086916 Human sp1
81	19	1.0	360	10	AD086916 Human sp1
82	19	1.0	360	10	AD086916 Human sp1
83	19	1.0	360	10	AD086916 Human sp1
84	19	1.0	360	10	AD086916 Human sp1
85	19	1.0	360	10	AD086916 Human sp1
86	19	1.0	360	10	AD086916 Human sp1
87	19	1.0	360	10	AD086916 Human sp1
88	19	1.0	360	10	AD086916 Human sp1
89	19	1.0	360	10	AD086916 Human sp1
90	19	1.0	360	10	AD086916 Human sp1
91	19	1.0	360	10	AD086916 Human sp1
92	19	1.0	360	10	AD086916 Human sp1
93	19	1.0	360	10	AD086916 Human sp1

94	19	1.0	579	13	AD055265	Adg55265 Novel can
95	19	1.0	606	5	AAH52281	Aah52281 Human AFP
96	19	1.0	735	13	AD512399	Ad512399 DNA encod
97	19	1.0	943	6	ABK70016	Abk70016 cDNA enco
98	19	1.0	943	9	ADA01379	Ada01379 Human PRO
99	19	1.0	943	9	ADA43808	Ada43808 Human cDN
100	19	1.0	943	9	ADA43576	Ada43576 Human cDN

ALIGNMENTS

```
RESULT 1
AAf27408
ID   AAF27408 standard; cDNA; 1883 BP.
XX
XX   AAF27408;
AC
XX   24-APR-2001 (first entry)
XX
XX   Human apoptosis-associated factor NT2RM1000558 cDNA, SEQ ID NO:3.
DE
XX
XX   Human: apoptosis-associated factor; NT2RM1000558; death effector domain;
KM   CED: caspase family cleavage domain; pro-apoptotic; drug screening;
KW   cell proliferation; ischaemic disease; chronic viral disease; ss.
XX
XX   Homo sapiens.
OS
XX   MO200104300-AL.
XX
XX   18-JUN-2001.
PD
XX
XX   06-JUL-2000; 2000WO-JP004516.
PF
XX
XX   08-JUL-1999; 99JP-00194179.
PR   18-OCT-1999; 99US-0159586P.
XX
XX   (HELI-) HELIX RES INST.
XX
XX   Ota T, Isegai T, Mshikawa T, Kawai Y, Miyoshi S, Satoh S;
PI   WPI: 2001-138348/14.
XX
XX   F-PSDB; AAB60387.
DR
XX
XX   Polynucleotide encoding an apoptosis-associated factor protein with death
PT   effector domain and caspase family-cleavage domain, useful in regulatig
PT   diseases with cell proliferation.
XX
XX   Claim 3; Page 44-47; 53pp; Japanese.
XX
XX   The invention relates to a novel human apoptosis-associated factor
CC   (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
CC   effector domain (DED) and a caspase family cleavage domain and is capable
CC   of inducing apoptosis in cells. The invention also relates to nucleic
CC   acids encoding the protein (AAF27407, AAF27408); variants of the protein
CC   (particularly dominant negative variants); vectors and host cells
CC   comprising a nucleic acid which encodes an apoptosis-associated factor
CC   of the invention; the recombinant production of the protein; an antibody
CC   against the protein; and methods of screening for compounds which can
CC   regulate apoptosis. The apoptosis-related factor is useful in regulating
CC   diseases associated with cell proliferation and in screening drug
CC   candidates e.g., for regulating cell proliferation or cell death in
CC   ischaemic diseases and chronic viral diseases. The present sequence
CC   represents a full-length cDNA encoding the human apoptosis-associated
CC   factor NT2RM1000558
XX
XX   Sequence 1883 BP; 343 A; 595 C; 588 G; 357 T; 0 U; 0 Other;
SQ
Query Match      100.0%; Score 1883; DB 5; Length 1883;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Db
1 AGCGCTAATATATGAGAAAGTGCACGAAAGATCCAAAACAAGTGGCTGGCGCCGTGCCC 60
QY
61 AGAAGTCATCGACCGCAGAAATCTGGCCGGATTCTGAGCTTTCCGCTCCCTCCCGG 120
Db
61 AGAAGTCATCGACCGCAGAAATCTGGCCGGATTCTGAGCTTTCCGCTCCCTCCCGG 120
QY
121 GGAATGCGCTATCCGGGTGACCCCGCGCTGTGGAGAGAGATAGTGGCTTGAC 180
Db
121 GGAATGCGCTATCCGGGTGACCCCGCGCTGTGGAGAGAGATAGTGGCTTGAC 180
QY
181 TACTACGGGATGCTGTGCTTACCGGTATGTTGAGAGTGTGGGGGCGCACTGACGAG 240
Db
181 TACTACGGGATGCTGTGCTTACCGGTATGTTGAGAGTGTGGGGGCGCACTGACGAG 240
QY
241 TCGAGCTGAGACTCTCTGCTTTCTGTGATGAGGCTCTTGCGCGCCGCGAGGCTTA 300
Db
241 TCGAGCTGAGACTCTCTGCTTTCTGTGATGAGGCTCTTGCGCGCCGCGAGGCTTA 300
QY
301 GCCCGGGCCCGACGGGCTTAAGCTCTGTGAGCTGAGAGCTGAGCGCGCGGCACTGGC 360
Db
301 GCCCGGGCCCGACGGGCTTAAGCTCTGTGAGCTGAGAGCTGAGCGCGCGGCACTGGC 360
QY
361 GAGAGCAACTGGGCTGTGGGGGCACTTGGCGGTGCTGGCCCGGCAAGCACTGCTG 420
Db
361 GAGAGCAACTGGGCTGTGGGGGCACTTGGCGGTGCTGGCCCGGCAAGCACTGCTG 420
QY
421 CCGCACTTGGCCCGCAGAGGGCGCGGCAAGTGTCTCCAGAAAGCTATATGAGCAC 480
Db
421 CCGCACTTGGCCCGCAGAGGGCGCGGCAAGTGTCTCCAGAAAGCTATATGAGCAC 480
QY
481 TCCAGCTCTTCAAGAAGACAGAGGGTAGTCCCGTCCGCTGGCGCACTGACAGATTCT 540
Db
481 TCCAGCTCTTCAAGAAGACAGAGGGTAGTCCCGTCCGCTGGCGCACTGACAGATTCT 540
QY
541 GCAAAATTCACAGAGGTACAGTGGGAGACAGAGGTCCCGCCCAACCAAGGGGACGGCGG 600
Db
541 GCAAAATTCACAGAGGTACAGTGGGAGACAGAGGTCCCGCCCAACCAAGGGGACGGCGG 600
QY
601 AGTGGGGCCCGGCCAGTGGTGTGTGCACAGCGCGCGGAGAGGGGCCCGACGGCACCC 660
Db
601 AGTGGGGCCCGGCCAGTGGTGTGTGCACAGCGCGCGGAGAGGGGCCCGACGGCACCC 660
QY
661 CAGCAGACGTCAGAGCCCGCCAGACTTCTCTGAAAGCAAAAGTGACTCTGTACATCCGG 720
Db
661 CAGCAGACGTCAGAGCCCGCCAGACTTCTCTGAAAGCAAAAGTGACTCTGTACATCCGG 720
QY
721 CTCGGGGTTCAGACAGAGTACTGCGAGCATGGGGCCAGCTTGGAGAGAGGGCGTGCATCC 780
Db
721 CTCGGGGTTCAGACAGAGTACTGCGAGCATGGGGCCAGCTTGGAGAGAGGGCGTGCATCC 780
QY
781 CGGCGGCCCGCAGGCGCTGGCGCGGCACTGAGAGTGTGGGCAAGGCCACCGAGTGTG 840
Db
781 CGGCGGCCCGCAGGCGCTGGCGCGGCACTGAGAGTGTGGGCAAGGCCACCGAGTGTG 840
QY
841 CGGTCAAGGAGACTGGGGCTGTGTGTTGTGACATCAATGTTCTCAGAGCTCTTCTATCTG 900
Db
841 CGGTCAAGGAGACTGGGGCTGTGTGTTGTGACATCAATGTTCTCAGAGCTCTTCTATCTG 900
QY
901 GAGCGCTTGTGGGGGACACTGAGTGGGCGCTGTGCTCAAGCCCTTGCGGGGCGTGTTC 960
Db
901 GAGCGCTTGTGGGGGACACTGAGTGGGCGCTGTGCTCAAGCCCTTGCGGGGCGTGTTC 960
QY
961 CTGACTGAGGCGCTGCGAGAGCTGTGGCGCGAGAGCTTGTGCGCTGTGTCACTGTG 1020
Db
961 CTGACTGAGGCGCTGCGAGAGCTGTGGCGCGAGAGCTTGTGCGCTGTGTCACTGTG 1020
QY
1021 GATGAGGCTGACTATGAGGCTGAGCGCGCGCTGTGCTGATGAGAGAGAGGGGGG 1080
Db
1021 GATGAGGCTGACTATGAGGCTGAGCGCGCGCTGTGCTGATGAGAGAGAGGGGGG 1080
QY
1081 CGGCGCGCGACAGAGGCTCTGATTCAGAGACTGGAGAGATTGATCCCACTCCAAAGTCT 1140
Db
1081 CGGCGCGCGACAGAGGCTCTGATTCAGAGACTGGAGAGATTGATCCCACTCCAAAGTCT 1140
```

D	1081	CGGGGCCGACAGAGGCGCTCTGATCCAGAGCACTGGCAGAGATTGATCCCACTCCAGTCT	1140
Q	1141	CCGGGCCACCTTCTTCCTGGAGGAGCAGCAATCTTACACCCCTTGAAGAGCCCTCCCAAGG	1200
D	1141	CCGGGCCACCTTCTTCCTGGAGGAGCAGCAATCTTACACCCCTTGAAGAGCCCTCCCAAGG	1200
Q	1201	ATGTGGCTCTGAGGGCTTAAACCATTTTCAGCTAGATTCTTCCAGACTCTCTTACC	1260
D	1201	ATGTGGCTCTGAGGGCTTAAACCATTTTCAGCTAGATTCTTCCAGACTCTCTTACC	1260
Q	1261	CCCAAGTGTGCCCCCTTAGCCTCCGGAGGGCGGGGGCTGGGCTGTATCTCAGAAAGGAGG	1320
D	1261	CCCAAGTGTGCCCCCTTAGCCTCCGGAGGGCGGGGGCTGGGCTGTATCTCAGAAAGGAGG	1320
Q	1321	GGCAGAGTACACACTACCAAAAGGCCCCCTGGACATTTGATCTGATCTTGGGGCTGT	1380
D	1321	GGCAGAGTACACACTACCAAAAGGCCCCCTGGACATTTGATCTGATCTTGGGGCTGT	1380
Q	1381	CTGCACTGTCCACAGGTGACACACTGCTCATGTCTCACACTGCCCTTGTGAGATCTTCC	1440
D	1381	CTGCACTGTCCACAGGTGACACACTGCTCATGTCTCACACTGCCCTTGTGAGATCTTCC	1440
Q	1441	CTGGGGCTCTGGCCCTTGGCTGTCCAGACACACATTTCTTGGCTTAAAGGCTTCTCTC	1500
D	1441	CTGGGGCTCTGGCCCTTGGCTGTCCAGACACACATTTCTTGGCTTAAAGGCTTCTCTC	1500
Q	1501	TCAGAGACCTCTAATTTTGACACAAACAACTGGGCTTCAGGCCATCATAGTGGGCACTGGA	1560
D	1501	TCAGAGACCTCTAATTTTGACACAAACAACTGGGCTTCAGGCCATCATAGTGGGCACTGGA	1560
Q	1561	GCTGGGGGTGCACATGGGGGCTGTCACTCTGCTCCACATCTCCAGCCAGCCAGGGGCTT	1620
D	1561	GCTGGGGGTGCACATGGGGGCTGTCACTCTGCTCCACATCTCCAGCCAGCCAGGGGCTT	1620
Q	1621	GCCGAGCTTCAATTTTACAGACTGATCTCTCTCACCTTCCCCCTGTGCTCAAGGCTGA	1680
D	1621	GCCGAGCTTCAATTTTACAGACTGATCTCTCTCACCTTCCCCCTGTGCTCAAGGCTGA	1680
Q	1681	ACATAGACTTGGACATTGGATGTCACTGGATGTTCACATGGAGTGTATTATGGCAGCATCA	1740
D	1681	ACATAGACTTGGACATTGGATGTCACTGGATGTTCACATGGAGTGTATTATGGCAGCATCA	1740
Q	1741	TACCAAGGCTCACTGTTTGCACATGGGGGCAAAACCACTGTAACAGCCACTTCTTGGAAAG	1800
D	1741	TACCAAGGCTCACTGTTTGCACATGGGGGCAAAACCACTGTAACAGCCACTTCTTGGAAAG	1800
Q	1801	GGAATGCAGAAAGGCTTGGGGGGTGCATGTGAGAAAGACCTTTTACAAATGATACCAATTAACT	1860
D	1801	GGAATGCAGAAAGGCTTGGGGGGTGCATGTGAGAAAGACCTTTTACAAATGATACCAATTAACT	1860
Q	1861	GCCTGTGAAAGGGCAATAGTGGG 1883	
D	1861	GCCTGTGAAAGGGCAATAGTGGG 1883	

RESULT 2
ABA94362
ID ABA94362 standard; DNA; 1230 BP.

XX ABA94362;
AC
XX
DT 26-MAR-2002 (first entry)
XX
XX Human APRG polypeptide (Incyte ID. 3102521CD1) encoding cDNA.
XX
XX APRG; apoptosis regulator; cytosolic; antiatherosclerotic; osteopathic;
KW antiatherosclerotic; hepatotropic; antipneumatic; antihelmintic; human;
KW antileukemic; antianemic; antischismatic; antihypoid; anti-HIV; cancer;
KW antiinflammatory; antidiabetic; antigrout; nephrotropic; ophthalmological;
KW immunosuppressive; dermatological; antileuc; antirheumatic; fungicide;
KW anticholeric; antibacterial; virucide; antiparasitic; protozoacide;
KW tranquilizer; vulnerary; gynecological; vasotropic; gene therapy; ss.

OS	Homo sapiens.
XX	
FH	Key
FT	127. .1107
FT	/tag= a
FT	/product= "APRG polypeptide"
XX	
PN	W0200192527-A2.
XX	
XX	
PD	06-DEC-2001.
XX	
XX	30-MAY-2001; 2001WO-US017581.
PR	01-JUN-2000; 2000US-0209407P.
XX	
PR	30-NOV-2000; 2000US-0250326P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Tang YT, Azimzai Y, Yue H, Buford N, Ding L, Elliott VS;
PI	Patterson C, Baughn MR;
XX	
DR	WPI; 2002-114350/15.
DR	P-PSDB; ABB07263.
XX	
PT	Novel human apoptosis regulator polypeptides and polynucleotides for
PT	diagnosing, preventing, treating cell proliferative, immunological and
PT	reproductive disorders and for identifying modulators of therapeutic use.
XX	
PS	Claim 5; Page 101; 103pp; English.
XX	
CC	The invention provides human apoptosis regulator (APRG) polypeptides and
CC	polynucleotides. The APRG polypeptides, polynucleotides and modulators
CC	are useful for diagnosis, treatment and prevention of cell proliferative,
CC	immunological and reproductive disorders. The cell proliferative
CC	disorders include cancers, actinic keratosis, arteriosclerosis,
CC	atherosclerosis, bursitis, cirrhosis, hepatitis, psoriasis, and
CC	immunological disorders include acquired immunodeficiency syndrome
CC	(AIDS), adult respiratory distress syndrome, Addison's disease,
CC	ankylosing spondylitis, amyloidosis, allergies, anemia, osteoporosis,
CC	autoimmune hemolytic anemia, asthma, autoimmune thyroiditis, Crohn's
CC	disease, contact dermatitis, diabetes mellitus, gout, Graves' disease,
CC	glomerulonephritis, rheumatoid arthritis, scleroderma, systemic lupus
CC	erythematosus, systemic sclerosis, ulcerative colitis, hemodialysis,
CC	uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
CC	infections and trauma. Reproductive disorders include disorders of
CC	prolactin production, infertility, endometriosis, polycystic ovary
CC	syndrome, ectopic pregnancies, galactorrhea, abnormal sperm physiology,
CC	disruptions of spermatogenesis, cancer of testis and prostate, impotence,
CC	carcinoma of male breast and gynecomastia. The APRG polynucleotides are
CC	useful for creating knockin humanized animals or transgenic animals to
CC	model human disease and to detect and quantify gene expression in
CC	biopsied tissues in which expression of APRG is correlated with disease.
CC	APRG, fragments of it and antibodies specific for APRG are useful as
CC	elements on a microarray which is useful to monitor or measure protein-
CC	protein interactions, drug-target interactions and gene expression
CC	profiles. The present sequence represents a human APRG polypeptide
CC	encoding cDNA
XX	
XX	
SQ	Sequence 1230 BP; 205 A; 382 C; 433 G; 210 T; 0 U; 0 Other;
	Query Match 60.0%; Score 1130; DB 6; Length 1230;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 AGGCGTATATATAGAGAGGTGCCAGAAAGATCCAAAACAAGTGGCTGCGGCGGTGCC
DB	4 AGGCGTATATATAGAGAGGTGCCAGAAAGATCCAAAACAAGTGGCTGCGGCGGTGCC
OY	61 AGAGTATCGAGAGCGCAGAAATCTGGCGGTTTGAGCTTGTTCGCTCCCTCCCGC
DB	64 AGAGTATCGAGAGCGCAGAAATCTGGCGGTTTGAGCTTGTTCGCTCCCTCCCGC
OY	121 GGAATGCGCTATTCGGGATGCAAGCCGCGCGCGCTGCTGGAGAGAGATGATGCTTGAC

```
Db 124 GGAATGGCCCTATCCGGGTGACCCCGGGCCCGTGTGGAGAGAGATAGTCCCTGGAC 183
Oy 181 TACTACGGGATGCTGTCGCTTCAACGATATGTTGAGAGTGTGGGCGGGCAATGACCGAG 240
Db 184 TACTACGGGATGCTGTCGCTTCAACGATATGTTGAGAGTGTGGGCGGGCAATGACCGAG 243
Oy 241 TCGAGCTGAGAGCTCTGAGCTTCTTGTGATGATGAGGCTCTGAGCGCCGCGAGGCTTA 300
Db 244 TCGAGCTGAGAGCTCTGAGCTTCTTGTGATGATGAGGCTCTGAGCGCCGCGAGGCTTA 303
Oy 301 GCGCGGGGCGGACGAGCGGCTTGAAGCTCTGAGAGTGTGGAGCGCGCGGCGAGTGGCG 360
Db 304 GCGCGGGGCGGACGAGCGGCTTGAAGCTCTGAGAGTGTGGAGCGCGCGGCGAGTGGCG 363
Oy 361 GAGAGCAACCTGAGGCTGTGGGCAACTCTGCGGTGTGGCCGCGGACGACTGCTG 420
Db 364 GAGAGCAACCTGAGGCTGTGGGCAACTCTGCGGTGTGGCCGCGGACGACTGCTG 423
Oy 421 CCGACCTGCGGCGCAAGCGGCGCGGCAAGTGTCTCAAGAACGCTATAGTATGGCAC 480
Db 424 CCGACCTGCGGCGCAAGCGGCGCGGCAAGTGTCTCAAGAACGCTATAGTATGGCAC 483
Oy 481 TCCAGCTCTTCAAAAGAGACAGAGGCTAGCTGCGGCTGCGGCTGCGGACGTCAGCA 540
Db 484 TCCAGCTCTTCAAAAGAGACAGAGGCTAGCTGCGGCTGCGGCTGCGGACGTCAGCA 543
Oy 541 GCAAAATTCTCAGAGGGTCACTGAGAGAGAGGCTCCCCCAACCAAGCGGCGAGCGG 600
Db 544 GCAAAATTCTCAGAGGGTCACTGAGAGAGAGGCTCCCCCAACCAAGCGGCGAGCGG 603
Oy 601 AGTCGGGCGCGGCGGAGTGTGTGTCAGACGCGGCGGAGAGAGGCGCGGCGGACCC 660
Db 604 AGTCGGGCGCGGCGGAGTGTGTGTCAGACGCGGCGGAGAGAGGCGCGGCGGACCC 663
Oy 661 CAGCAGCACTCAGAGCGCCCGCAGACCTTCTCTGAAGGCAAACTGACCTGTACATCCG 720
Db 664 CAGCAGCACTCAGAGCGCCCGCAGACCTTCTCTGAAGGCAAACTGACCTGTACATCCG 723
Oy 721 CTCGGGCTTCAAGAGTACTGCGAGCAATGAGGCGAGCTTGGAGAGAGGCGGAGTCC 780
Db 724 CTCGGGCTTCAAGAGTACTGCGAGCAATGAGGCGAGCTTGGAGAGAGGCGGAGTCC 783
Oy 781 CCGCGGCGCCGAGCGCTGTGGCGCGGAGCTGAGAGTGTGTGGGCGGACCGAGTGTG 840
Db 784 CCGCGGCGCCGAGCGCTGTGGCGCGGAGCTGAGAGTGTGTGGGCGGACCGAGTGTG 843
Oy 841 CGCTCAAGGAGCCTGAGGCTGTGTGTGTGTGATCAAGTTCCTCAAGCTCTCTATCTG 900
Db 844 CGCTCAAGGAGCCTGAGGCTGTGTGTGTGTGATCAAGTTCCTCAAGCTCTCTATCTG 903
Oy 901 GAGGCTTGTGGGCGAGTACTGAGTGTGCGCTCTGACAGGCGCTGCGGGCGTGTTC 960
Db 904 GAGGCTTGTGGGCGAGTACTGAGTGTGCGCTCTGACAGGCGCTGCGGGCGTGTTC 963
Oy 961 CTGACTGAGGCGCTGTGAGAGGCTGTGGCGGAGGCTGTGTGCTGTGTGAGTGTG 1020
Db 964 CTGACTGAGGCGCTGTGAGAGGCTGTGGCGGAGGCTGTGTGCTGTGTGAGTGTG 1023
Oy 1021 GATGAGGCTGATATAGGCTGGCGCGGCTTGTGTGATGAGAGAGAGAGGAGG 1080
Db 1024 GATGAGGCTGATATAGGCTGGCGCGGCTTGTGTGATGAGAGAGAGAGGAGG 1083
Oy 1081 CCGCGGCGGACAGAGGCTCTGATCAAGAGTGTGAGAGATTTGATCCACCTCAAGTCT 1140
Db 1084 CCGCGGCGGACAGAGGCTCTGATCAAGAGTGTGAGAGATTTGATCCACCTCAAGTCT 1143
Oy 1141 CCGGCGGCACTTCTCTGTGAGAGACCAATCTTACCCCT 1181
Db 1144 CCGGCGGCACTTCTCTGTGAGAGACCAATCTTACCCCT 1184
```

, RESULT 3

```
AD086891
ID AD086891 standard; cDNA; 1979 BP.
XX
AC AD086891;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #3766.
XX
KW human; tumour-associated antigenic target; TAT; cytosolic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH ) GENENTECH INC.
XX (MUTD ) KU T D.
XX (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-53430/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 3766; 5504bp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
```



```

QY 26 GAAAGATCCAAAACAAGTGGCTGCGGCGCTGCGCCAGAGACTCATCGACGCCAGAACTTG 85
DB 23 GAAAGATCCAAAACAAGTGGCTGCGGCGCTGCGCCAGAGACTCATCGACGCCAGAACTTG 82
QY 86 GCGGGGTTCTGAGACTTGTTCCTGCTCCCTCCCGGGAAATGGGCTATCCGGGTCCAGCC 145
DB 83 GCGGGGTTCTGAGACTTGTTCCTGCTCCCTCCCGGGAAATGGGCTATCCGGGTCCAGCC 142
QY 146 CCGCCCCGCTGCTGAGAGAGAGATGAGTGCCTGAGACTACTAGCGGAGTGTCTGCTTCAAC 205
DB 143 CCGCCCCGCTGCTGAGAGAGAGATGAGTGCCTGAGACTACTAGCGGAGTGTCTGCTTCAAC 202
QY 206 GATATGTTGAGAGTGTGGGCGGCAACTGACCGAGTGCAGAGCTTGAGCTTCTGAGCTTTT 265
DB 203 GATATGTTGAGAGTGTGGGCGGCAACTGACCGAGTGCAGAGCTTGAGCTTCTGAGCTTTT 262
QY 266 TGTGTGATGAGGCTCTCTGGGCGCGCGAGAGCTTACCCCGGGCCCGCAGAGGCTTAGAG 325
DB 263 TGTGTGATGAGGCTCTCTGGGCGCGCGAGAGCTTACCCCGGGCCCGCAGAGGCTTAGAG 322
QY 326 TCTGTGTGAGAGTGCAGCGCGCGGCGAGTGCAGCGAGCAACTGCGGCTGCTGAGGCG 385
DB 323 TCTGTGTGAGAGTGCAGCGCGCGGCGAGTGCAGCGAGCAACTGCGGCTGCTGAGGCG 382
QY 386 AACTCTGCGCGTGTGCGCCGCGCACGACTGTGCGCACTGCGCGCGCAAGCGGCGCGC 445
DB 383 AACTCTGCGCGTGTGCGCCGCGCACGACTGTGCGCACTGCGCGCGCAAGCGGCGCGC 442
QY 446 GCGCAATGTCTCCAGAAACCTATAGCTATAGGACCTTCCAGCTTTTAAAGAGAGACAGAG 505
DB 443 GCGCAATGTCTCCAGAAACCTATAGCTATAGGACCTTCCAGCTTTTAAAGAGAGACAGAG 502
QY 506 GTAGCTGCGCTGCGCTGCGCACTCAAGCAGTTCGCAAAATTTCTAGACAGAGGTCACTGAG 565
DB 503 GTAGCTGCGCTGCGCTGCGCACTCAAGCAGTTCGCAAAATTTCTAGACAGAGGTCACTGAG 562
QY 566 AGACAGGCTTCCCCCAACCAAGCGGCGAGTGCAGGCTGCGGCGCGCCAGTGTGAGTG 625
DB 563 AGACAGGCTTCCCCCAACCAAGCGGCGAGTGCAGGCTGCGGCGCGCCAGTGTGAGTG 622
QY 626 CCGAGAGCGGCGGAGAGAGGCGCCCGACCGGCAACCCAGAGCGAGTCAAGAGCCCGCAGAG 685
DB 623 CCGAGAGCGGCGGAGAGAGGCGCCCGACCGGCAACCCAGAGCGAGTCAAGAGCCCGCAGAG 682
QY 686 CTTCCTCTGAGAGCAAAAGTGAACCTGTGACATCCGCGCTCGGCTTCAGAGCAGATCTGCG 745
DB 683 CTTCCTCTGAGAGCAAAAGTGAACCTGTGACATCCGCGCTCGGCTTCAGAGCAGATCTGCG 742
QY 746 AGCATGCGGCGAGCTTTGAGAGCAGGCGTGGCATCCCGCGCGCCCAAGGCGTGGCGGCG 805
DB 743 AGCATGCGGCGAGCTTTGAGAGCAGGCGTGGCATCCCGCGCGCCCAAGGCGTGGCGGCG 802
QY 806 AGCTGAGACGTGTTTGGGCGAGGCCACCGGAGTGTGGCTTCAAGGAGCTTGGGCTTGTG 865
DB 803 AGCTGAGACGTGTTTGGGCGAGGCCACCGGAGTGTGGCTTCAAGGAGCTTGGGCTTGTG 862
QY 866 TTGTGATCAATCAAGTTCCTGAGAGCTCTCTATCTGAGCGCTTCTGAGGCGCATCTGAG 925
DB 863 TTGTGATCAATCAAGTTCCTGAGAGCTCTCTATCTGAGCGCTTCTGAGGCGCATCTGAG 922
QY 926 GTGGCGCGCTGCTGAGAGCGCTTGGGCGGCGTGTCTTCACTGAGAGCCCTTGCAGAGGCTG 985
DB 923 GTGGCGCGCTGCTGAGAGCGCTTGGGCGGCGTGTCTTCACTGAGAGCCCTTGCAGAGGCTG 982
QY 986 TGGGCGCGGAGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045
DB 983 TGGGCGCGGAGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
QY 1046 GCGCGCGCGCTGTTGCTGATGAGAGAGAGAGGCGGCGCGCCCGAGAGAGAGGCTCTCTGAT 1105
DB 1043 GCGCGCGCGCTGTTGCTGATGAGAGAGAGAGGCGGCGCGCCCGAGAGAGAGGCTCTCTGAT 1102

```

```

QY 1106 CCGAGACTGGCAGAGATTGATCCCAAGTCTTCGCGGCGCACTTCTCTGGAGAGAC 1165
DB 1103 CCGAGACTGGCAGAGATTGATCCCAAGTCTTCGCGGCGCACTTCTCTGGAGAGAC 1162
QY 1166 GACCATCTACCCCT 1181
DB 1163 GACCATCTACCCCT 1178

RESULT 5
AAH33283
ID AAH33283 standard; cDNA; 2045 BP.
XX
AC AAH33283;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:339.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN MO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000MO-US026524.
XX
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
XX
DR P-PSDB; AAG73852.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 1; Page 2451; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where the
XX proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytostatic activity and can be used in gene therapy
XX and vaccine production. N and P may be used in the prevention, diagnosis
XX and treatment of diseases associated with inappropriate P expression. For
XX example, N and P may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of P by expressing inactive proteins or to
XX supplement the patient's own production of P. Additionally, N may be used
XX to produce the colon cancer-associated Ps, by inserting the nucleic acids
XX into a host cell and culturing the cell to express the proteins. N and P
XX can be used in the prevention, diagnosis and treatment of colorectal
XX carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
XX sequences used in the exemplification of the present invention. N.B.
XX Pages 666 to 682 and page 7053 of the sequence listing were missing at
XX time of publication, meaning no sequences are present for SEQ ID NO:1027
XX to 1052, 7921 and 7922
XX
SO Sequence 2045 BP; 457 A; 620 C; 604 G; 364 T; 0 U; 0 Other;

Query Match 58.7%; Score 1105; DB 4; Length 2045;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 26 GAAAGATCCAAAACAAGTGGCTGCGGCGCTGCGCCAGAGACTCATCGACGCCAGAACTTG 85
DB 23 GAAAGATCCAAAACAAGTGGCTGCGGCGCTGCGCCAGAGACTCATCGACGCCAGAACTTG 82

```


QY 86 GCCGGGTTCTGAGCTTGTTCGGCTCCCTCCCGGAGATGCGCTATCCGGGTGACCC 145
 DB 83 GCCGGGTTCTGAGCTTGTTCGGCTCCCTCCCGGAGATGCGCTATCCGGGTGACCC 142
 QY 146 CGGCCCCGTCGTGGAGAGAGATGAGTCTGGAATACTAAGGAGTCTGCTTACC 205
 DB 143 CGGCCCCGTCGTGGAGAGAGATGAGTCTGGAATACTAAGGAGTCTGCTTACC 202
 QY 206 GTATGTTCCAGAGTGTGGGGGCACTGACCAAGTGGAGTGGAGTCTGCTTACC 265
 DB 203 GTATGTTCCAGAGTGTGGGGGCACTGACCAAGTGGAGTGGAGTCTGCTTACC 262
 QY 266 TGTGTGATGAGGCTCTGCGCGCCGCGAGAGCTTACGCGGCGCGAGGCTTACAGC 325
 DB 263 TGTGTGATGAGGCTCTGCGCGCCGCGAGAGCTTACGCGGCGCGAGGCTTACAGC 322
 QY 326 TCTGTGTGAGTGTGAGAGCGCGCGGAGTGTGCGGAGAGCAACTGCGGCTGTGGGGC 385
 DB 323 TCTGTGTGAGTGTGAGAGCGCGCGGAGTGTGCGGAGAGCAACTGCGGCTGTGGGGC 382
 QY 386 AACTCTGCGCGTGTGGCGCGCCGAGCACTGCTGCGGCACTGGCGCGGAGAGGCGGC 445
 DB 383 AACTCTGCGCGTGTGGCGCGCCGAGCACTGCTGCGGCACTGGCGCGGAGAGGCGGC 442
 QY 446 GGCAGATGTCTCCAGAAAGCTATAGTATGAGCACTTCAGACTCTTCAAAGAGAGAGAGG 505
 DB 443 GGCAGATGTCTCCAGAAAGCTATAGTATGAGCACTTCAGACTCTTCAAAGAGAGAGAGG 502
 QY 506 GTAGCTGCGTGTGCGCGTGTGAGCAAGTGTGCAAAATTTCTGAGCAGAGGTCAGTGGG 565
 DB 503 GTAGCTGCGTGTGCGCGTGTGAGCAAGTGTGCAAAATTTCTGAGCAGAGGTCAGTGGG 562
 QY 566 AGACAGGCTCTCCCGCAACCAAGCGGAGAGGCGGAGTCCGGGCGCGGCGCAAGTGTGGTGG 625
 DB 563 AGACAGGCTCTCCCGCAACCAAGCGGAGAGGCGGAGTCCGGGCGCGGCGCAAGTGTGGTGG 622
 QY 626 CCAGAGCGCGCGGAGAGAGGCGCGCAGCGCAGCACTTCAGAGTGTGAGAGCGCGCAGAC 685
 DB 623 CCAGAGCGCGCGGAGAGAGGCGCGCAGCGCAGCACTTCAGAGTGTGAGAGCGCGCAGAC 682
 QY 686 CTTCCTCTGAGAGCAAAAGTACCTGTGATCCGGCTCCGGGTTTCCAGAGAGTACTGGC 745
 DB 683 CTTCCTCTGAGAGCAAAAGTACCTGTGATCCGGCTCCGGGTTTCCAGAGAGTACTGGC 742
 QY 746 AGATGAGCGCGCTTGTGAGAGAGGCGTGTGATCCGGCGCGGCGCGCAGCGCTGTGGCGGC 805
 DB 743 AGATGAGCGCGCTTGTGAGAGAGGCGTGTGATCCGGCGCGGCGCGCAGCGCTGTGGCGGC 802
 QY 806 AGCTGAGCGTGTGAGAGAGGCGCGCAGCGTGTGATCCGGCGCGCTGTGGC 865
 DB 803 AGCTGAGCGTGTGAGAGAGGCGCGCAGCGTGTGATCCGGCGCGCTGTGGC 862
 QY 866 TTTGTGATCAAGTGTCTGAGAGCTCTCTATCTGAGAGCGCTTCTGGGGGCACTACCTGA 925
 DB 863 TTTGTGATCAAGTGTCTGAGAGCTCTCTATCTGAGAGCGCTTCTGGGGGCACTACCTGA 922
 QY 926 GTGGCGCGCTGTGAGAGAGGCGCGCTGTGATCCGGCGCGCTGTGGAGAGGCTGTG 985
 DB 923 GTGGCGCGCTGTGAGAGAGGCGCGCTGTGATCCGGCGCGCTGTGGAGAGGCTGTG 982
 QY 986 TGGGCGCGAGGCTGTGCGCTGTGATCCGGCGCGCTGTGATCCGGCGCGCTGTGGC 1045
 DB 983 TGGGCGCGAGGCTGTGCGCTGTGATCCGGCGCGCTGTGATCCGGCGCGCTGTGGC 1042
 QY 1046 GCGCGCGCTGTGATCCGGCGCGCTGTGATCCGGCGCGCTGTGGAGAGGCTGTGAT 1105
 DB 1043 GCGCGCGCTGTGATCCGGCGCGCTGTGATCCGGCGCGCTGTGGAGAGGCTGTGAT 1102
 QY 1106 CCAGAGCTGTGAGAGTGTGATCCGGCGCGCTGTGATCCGGCGCGCTGTGGAGAGGCT 1165
 DB 1103 CCAGAGCTGTGAGAGTGTGATCCGGCGCGCTGTGATCCGGCGCGCTGTGGAGAGGCT 1162

QY 1166 GACCATCTTACCCCT 1181
 DB 1163 GACCATCTTACCCCT 1178
 RESULT 6
 AAD40080
 ID AAD40080 standard; DNA; 1924 BP.
 XX
 AC AAD40080;
 XX
 DT 22-OCT-2002 (first entry)
 XX
 DE Human DED4 (death effector domain) gene.
 XX
 KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
 KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
 KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
 KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
 KW immunosuppressive; gene therapy; antisense therapy; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS
 FT 91..1044
 FT /tag= a
 FT /product= "Human DED4"
 FT /note= "No stop codon"
 FT /partial
 FT misc_feature
 FT 157..222
 FT /*tag= b
 FT /note= "Nuclear localisation sequence"
 XX
 PN M0200240680-A2.
 XX
 PD 23-MAY-2002.
 XX
 PE 15-NOV-2001; 2001MO-US044844.
 XX
 PR 17-NOV-2000; 2000US-00715893.
 PR 29-JUN-2001; 2001US-0301889P.
 XX
 PA (BURN-) BURHAM INST.
 XX
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Stemmer-Wiemer F;
 XX
 DR WPI; 2002-500222/53.
 DR P-PSDB; AAE24860.
 XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
 XX
 PS Claim 19; Page 184-186; 209pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising a death
 CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
 CC is useful for identifying a binding agent, preferably a protein or a drug
 CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
 CC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DD protein), DED4 or
 CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
 CC detecting the association of the domain and the candidate binding agent,
 CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
 CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
 CC spectroscopy (MS) and FPA. The invention is useful for modulating the
 CC level of a cell process such as cell proliferation, cell adhesion, cell
 CC stress responses, responses to microbial infection and B cell
 CC immunoglobulin class switching, in particular apoptosis within a cell.
 CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
 CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
 CC CTDD DD protein is useful for detecting a Chlamydia infection. The
 CC invention is useful for modulating the activity of oncogenic proteins,

for treating a pathology caused by the oncogenic proteins and for treating bacterial infections by modulating the activity of bacterial proteins. The protein and antibody specific for it are useful for discovery of drugs that suppress infection, inflammation, allergy, sepsis, autoimmunity, allograft rejection and other diseases. The protein is useful for treating immune-based pathologies, pathologies associated with cell division, inflammatory diseases such as sepsis, fibrosis, arthritis, graft versus host disease. The invention is used in anticancer therapy and gene therapy. The present sequence is human DED4 gene

Sequence 1924 BP; 360 A; 599 C; 594 G; 368 T; 0 U; 3 Other;

Query Match 57.2%; Score 1078; DB 6; Length 1924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1128; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

CC CGTCCGCGAGAGTACTGCGAGCGCAAGATCTGAGCGGGTTCTGAGCTTTCGCGCTCC 112
DB CCGTCCGCGAGAGTACTGCGAGCGCGAAGATCTGAGCGGGTTCTGAGCTTTCGCGCTCC 79
QY 113 CTCGCCCGGGAATGCGGCTATCCGGGTGCAACCCCGGCCCTGCTGGAGAGAGATGAGT 172
DB CTCGCCCGGGAATGCGGCTATCCGGGTGCAACCCCGGCCCTGCTGGAGAGAGATGAGT 139
QY 173 GCCTGACACTACTACGGGAGTCTGCTGCTTACCGTATGTTTCAAGGTGTGGGGGCAAC 232
DB 140 GCCTGACACTACTACGGGAGTCTGCTGCTTACCGTATGTTTCAAGGTGTGGGGGCAAC 199
QY 233 TGACGAGTGCAGTGCAGTGCAGTCTCTGAGGCTCTGCTGAGTGCAGTGCAGTGCAGT 292
DB 200 TGACGAGTGCAGTGCAGTGCAGTCTCTGAGGCTCTGCTGAGTGCAGTGCAGTGCAGT 259
QY 293 GAGGCTTACGCGCGGCGCGGAGCGGCTGAGGCTCTGCTGAGTGCAGTGCAGTGCAGT 352
DB 260 GAGGCTTACGCGCGGCGCGGAGCGGCTGAGGCTCTGCTGAGTGCAGTGCAGTGCAGT 319
QY 353 AGTGCAGGCGAGAGCACTGCGGCTGCTGAGGCACTCTCTGCGGCTGCTGCGGCGCAAG 412
DB 320 AGTGCAGGCGAGAGCACTGCGGCTGCTGAGGCACTCTCTGCGGCTGCTGCGGCGCAAG 379
QY 413 ACCTGCTGCAGCACTGCGGCGCGAGCGGCGCGGCAAGTGTCTCCAGAAAGCTATAGCT 472
DB 380 ACCTGCTGCAGCACTGCGGCGCGAGCGGCGCGGCAAGTGTCTCCAGAAAGCTATAGCT 439
QY 473 ATGGCACTTTCAGCTTCTTCAAGAGAGACAGAGGGTATGCTGCGCGCTGCGCACTCAA 532
DB 440 ATGGCACTTTCAGCTTCTTCAAGAGAGACAGAGGGTATGCTGCGCGCTGCGCACTCAA 499
QY 533 GCACTTCTGCAAAATTTCTCAGCAGAGGTCAATGAGGAGAGAGAGAGAGAGAGAGAG 592
DB 500 GCACTTCTGCAAAATTTCTCAGCAGAGGTCAATGAGGAGAGAGAGAGAGAGAGAGAG 559
QY 593 AGCGGCGAGAGTCCGGGCGCGGCGCGAGTGTGTCAGAGAGCGGCGGAGAGAGAGAGAG 652
DB 560 AGCGGCGAGAGTCCGGGCGCGGCGCGAGTGTGTCAGAGAGCGGCGGAGAGAGAGAGAG 619
QY 653 CCGCAGCCCGAGCAGCAGTCAAGAGCCCGCAGACCTTCTTGAAGCAAAAGTGAAGCTGTG 712
DB 620 CCGCAGCCCGAGCAGCAGTCAAGAGCCCGCAGACCTTCTTGAAGCAAAAGTGAAGCTGTG 679
QY 713 AACTCCGGGCTCCGGGTTTCAGAGAGAGTATGCGAGAGATGGGCCAGCTTGTGAGAGAGGG 772
DB 680 AACTCCGGGCTCCGGGTTTCAGAGAGAGTATGCGAGAGATGGGCCAGCTTGTGAGAGAGGG 739
QY 773 TGGCATCCCGGCGGCGCGGCGCTGCGCGGCGAGCTGAGCGTGTGTTGGGCAAGCCACCG 832
DB 740 TGGCATCCCGGCGGCGCGGCGGCGCTGCGCGGCGAGCTGAGCGTGTGTTGGGCAAGCCACCG 799
QY 833 CAGTGTGCGCTCAAGGAGCTTGTGAGTGTGTTGTGACATCAAGTTCTCAGAGCTCT 892
DB 800 CAGTGTGCGCTCAAGGAGCTTGTGAGTGTGTTGTGACATCAAGTTCTCAGAGCTCT 859
QY 893 CCTATCTGAGAGCGCTTCTGTGGGCGAGCTACTAGTGGCGCCCTGCTGCGAGCGCCCTGCGGG 952

DB 860 CCTATCTGAGAGCGCTTCTGTGGGCGAGTACTAGTGGCGCCCTGCTGCGAGCGCTCGGG 919
QY 953 GCGTGTCTCTGACTGAGGCGCTTCCGAGAGAGCTGTGGGCGGAGAGCTGTTCGCTCTGG 1012
DB 920 GCGTGTCTCTGACTGAGGCGCTTCCGAGAGAGCTGTGGGCGGAGAGCTGTTCGCTCTGG 979
QY 1013 TCAGTGTGATGAGGCTGACTATGAGGCTGCGCGCGCGCTGTGCTGATGAGAGAGG 1072
DB 980 TCAGTGTGATGAGGCTGACTATGAGGCTGCGCGCGCGCTGTGCTGATGAGAGAGG 1039
QY 1073 AAGGGGCGGCGCGCGCGAGAGAGGCGCTCTGATCCAGAGCTGGCAGATTGATCCACCT 1132
DB 1040 AAGGGGCGGCGCGCGCGAGAGAGGCGCTCTGATCCAGAGCTGGCAGATTGATCCACCT 1099
QY 1133 CCAAGTCTCCGGGCGCACCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1181
DB 1100 CCAAGTCTCCGGGCGCACCTTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1148

RESULT 7
AAD59062
ID AAD59062 standard; cDNA; 1924 BP.
XX
AC AAD59062;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human DED4 full length gene.
XX
KW Human; death Domain; DD; death effector domain; DED; cell proliferation;
KW Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;
KW neural growth factor receptor-interacting death domain; cell adhesion;
KW vasotropic; microbial infection; inflammation; allograft rejection; CTDD;
KW cell stress response; benign prostatic hyper trophy; antibacterial; NMD;
KW apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;
KW neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;
KW keloid; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 91..1044
FT /tag= a
FT /product= "Human DED4 protein"
FT /note= "No stop codon"
FT /partial
FT 157..222
FT /tag= b
FT /note= "Nuclear localisation sequence"
XX
PN US2003049702-A1.
XX
PD 13-MAR-2003.
XX
PF 15-NOV-2001; 2001US-00001254.
XX
PR 17-NOV-2000; 2000US-00715893.
PR 17-NOV-2000; 2000US-0367360P.
PR 29-JUN-2001; 2001US-0301889P.
XX
PA (REED/) REED J C.
PA (GODZ/) GODZIK A.
PA (PAWL/) PAWLOWSKI K.
PA (PIOR/) FIORENTINO L.
PA (LEES/) LEE S H.
PA (ROTH/) ROTH W.
PA (STEN/) STENNER-LIEWEN F.
XX
XX Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W,
PI Stenner-Liewen F;
XX
DR MPI, 2002-500222/53.

DR WPI; 2000-587514/55.
 XX P-PSDB; AAB58420.
 XX Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer.

PS Claim 1; Page 772-773; 1425pp; English.

XX Polynucleotide sequences AAF1982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548, lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive; and
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotoxic; anti-infective; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the protein
 CC or polynucleotide sequences. The lung cancer associated polynucleotide
 CC sequences may be used for detection of lung cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The proteins may be used to treat disorders such as
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
 CC cardiovascular, renal, and proliferative disorders. The proteins may also
 CC be used in the treatment of wounds and infectious diseases.
 CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
 CC used in the course of the invention for the identification and
 CC characterisation of the polynucleotide and protein sequences

SQ Sequence 2044 BP; 457 A; 620 C; 603 G; 364 T; 0 U; 0 Other;

Query Match 52.3%; Score 984; DB 3; Length 2044;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1154; Conservativity 0; Mismatches 1; Indels 1; Gaps 1;

QY 26 GAAAGATCCAAAACAAGTGGCTGGCGCGCTGCGCCAGAGTCAATCGAGCCGAAATCTG 85
 Db 23 GAAAGATCCAAAACAAGTGGCTGGCGCGCTGCGCCAGAGTCAATCGAGCCGAAATCTG 82
 QY 86 GCCGGGTTTGAGCTTGTTCCTGCTCCCTCCCGGGAATGCGCTATCCGGGTGACCC 145
 Db 83 GCCGGGTTTGAGCTTGTTCCTGCTCCCTCCCGGGAATGCGCTATCCGGGTGACCC 142
 QY 146 CGGCCCCGCTGGGAGAGAGATGCTGAGCTAATCAACGAGAGCTGCTGCTTACC 205
 Db 143 CGGCCCCGCTGGGAGAGAGATGCTGAGCTAATCAACGAGAGCTGCTGCTTACC 202,
 QY 206 GTATGTTCCAGAGTGTGGCGCGGCACTGACCGAGTCCGAGCTTCTGCGCTTTC 265
 Db 203 GTATGTTCCAGAGTGTGGCGCGGCACTGACCGAGTCCGAGCTTCTGCGCTTTC 262
 QY 266 TGTGTGATGAGGCTCTTGGCGCGCGGAGAGCTTAAAGCCCGGCGGCGCTTAAGC 325
 Db 263 TGTGTGATGAGGCTCTTGGCGCGCGGAGAGCTTAAAGCCCGGCGGCGCTTAAGC 322
 QY 326 TCTGTGTGAGAGTGGAGCGCGCGGAGAGTGGCGGCGGAGAGCACTGCGGCTGCGGCG 385
 Db 323 TCTGTGTGAGAGTGGAGCGCGCGGAGAGTGGCGGCGGAGAGCACTGCGGCTGCGGCG 382
 QY 386 AACTCTGCGCGTGTGCGCGCGGCACTGCTGCGGCACTTCTGAGCGCGGCAAGCGCGC 445
 Db 383 AACTCTGCGCGTGTGCGCGCGGCACTGCTGCGGCACTTCTGAGCGCGGCAAGCGCGC 442
 QY 446 GGCAGTGTCTCCAGAAAGCTATAGTATGCACTTCCAGCTTCTTAAAGAGAGACAGAG 505
 Db 443 GGCAGTGTCTCCAGAAAGCTATAGTATGCACTTCCAGCTTCTTAAAGAGAGACAGAG 502
 QY 506 GTAGCTGCGCGTGGCGCGGCACTGAGAGCTTCAAAATTTCTGAGCGAGGCTCAATGGG 565
 Db 503 GTAGCTGCGCGTGGCGCGGCACTGAGAGCTTCAAAATTTCTGAGCGAGGCTCAATGGG 562
 QY 566 AGACAGAGCTCCCGCAACCAAGCGGAGCGGCGAGTCCGGGCGCGGCGGAGGTGTG 625
 Db 563 AGACAGAGCTCCCGCAACCAAGCGGAGCGGCGAGTCCGGGCGCGGCGGAGGTGTG 622

QY 626 CCAAGACGGCGGAGAGAGGGGCCCCAGCCGACACCCAGACAGTACAGAGCCCGGACAGC 685
 Db 623 CCAGACGGCGGAGAGAGGGGCCCCAGCCGACACCCAGACAGTACAGAGCCCGGACAGC 682
 QY 686 CTTCCCTGAGGCAAAAGTGAATGATCATCCGGCTCCGGGTTCCAGAGAGTACTGCG 745
 Db 683 CTTCCCTGAGGCAAAAGTGAATGATCATCCGGCTCCGGGTTCCAGAGAGTACTGCG 742
 QY 746 AGCATGGGCGAGCCTTGGAGCAGGCGGCTGCAATCCCGGCGGCGGCGGCGGCGGCG 805
 Db 743 AGCATGGGCGAGCCTTGGAGCAGGCGGCTGCAATCCCGGCGGCGGCGGCGGCGGCGG 802
 QY 806 AGCTGAGCTGTTTGGGCGAGGCGCACCGGAGTGTGCTCAAGAGACCTGGGCTCTGTG 865
 Db 803 AGCTGAGCTGTTTGGGCGAGGCGCACCGGAGTGTGCTCAAGAGACCTGGGCTCTGTG 862
 QY 866 TTTGTGACATCAAGTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGCGGAGTACTGA 925
 Db 863 TTTGTGACATCAAGTCTCAGAGCTCTCTATCTGAGCGCTTCTGGGCGGAGTACTGA 922
 QY 926 GTGGCGCCCTGTGTCAGAGGCGCTGCGGCGGCGGCTTCTGACTGAGGCGCTGAGAGGCTG 985
 Db 923 GTGGCGCCCTGTGTCAGAGGCGCTGCGGCGGCGGCTTCTGACTGAGGCGCTGAGAGGCTG 981
 QY 986 TGGGCGGCGAGGCTGTTCCCTGCTGTGATGATGATGATGATGATGATGATGATGATG 1045
 Db 982 TGGGCGGCGAGGCTGTTCCCTGCTGTGATGATGATGATGATGATGATGATGATGATG 1041
 QY 1046 GGGCGCCCTGTGTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1105
 Db 1042 GGGCGCCCTGTGTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1101
 QY 1106 CCAAGATGCGAGAGATTTGATCCACTCCAGTCTTCCGCGGCACTTCTCTGAGAGAC 1165
 Db 1102 CCAAGATGCGAGAGATTTGATCCACTCCAGTCTTCCGCGGCACTTCTCTGAGAGAC 1161
 QY 1166 GACCATCTTACCCCT 1181
 Db 1162 GACCATCTTACCCCT 1177

RESULT 9
 ADCT9259
 ID ADCT9259 standard; cDNA; 981 BP.
 XX
 AC ADCT9259;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human DBDD2 encoding cDNA SEQ ID NO:1.
 XX
 KW human; death effector domains containing DNA-binding protein;
 KW DBD-containing DNA-binding protein; DBDD2; cell death; gene therapy;
 KW cytoskeletal; cancer; chronic myeloid leukemia; gene; B.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..981
 FT /*tag= a
 FT /product= "DBDD2"
 PN MO2003054195-A1.
 PD 03-JUL-2003.
 XX
 PF 20-DEC-2002; 2002MO-JP013371.
 XX
 PR 20-DEC-2001; 2001JP-00387854.
 PR 18-JUL-2002; 2002JP-00209458.
 XX
 PA (MORG) MORINAGA MILK IND CO LTD.
 XX

PI Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
 XX WPI; 2003-569246/53.
 DR P-PSDB; ADB79260.
 XX
 PT DNA encoding cell death proteins for treatment of kidney, large intestine
 and prostate cancers and leukemia.
 XX
 PS Claim 2; Page 18-20; 26pp; Japanese.
 XX
 CC The present sequence encodes a human death effector domain (DED)
 containing DNA-binding protein (DED) protein, designated DED2, that
 causes cell death. Also described: (1) primer and probe for investigation
 of the DED2 gene; and (2) reagents for gene therapy. DED2 has
 CC cytotactic activity. DED2 can be used in the diagnosis and treatment of
 CC cancers of the kidney, large intestine and prostate, and acute and
 CC chronic myeloid leukaemia.
 XX
 SQ Sequence 981 BP; 151 A; 303 C; 363 G; 164 T; 0 U; 0 Other;

Query Match 49.4%; Score 930; DB 10; Length 981;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 980; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 124 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGAGATGCTCTGACTAC 183
 DB 1 ATGGCGCTATCCGGGTGACCCCGCCCGTGTGGAGAGAGATGCTCTGACTAC 60
 QY 184 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGGCACTGACCGAGTGC 243
 DB 61 TACGGATGCTGTGCTTCAACCGTATGTTGAGGTGTGGCGGGCACTGACCGAGTGC 120
 QY 244 GAGCTGAGCTCTGCGCTTCTGCTGATGAGAGCTCTGCGCGCGCGGAGGCTTAC 303
 DB 121 GAGCTGAGCTCTGCGCTTCTGCTGATGAGAGCTCTGCGCGCGCGGAGGCTTAC 180
 QY 304 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGGAGCTGCGGAG 363
 DB 181 CGGGCCCGCAGCGGCTAGAGCTCTGCTGAGCTGAGCGCGCGGAGCTGCGGAG 240
 QY 364 AGCAACCTGGGCTGCTGGGCACTCTCGCGGTGCTGGCCCGCAGACTGCTGCGG 423
 DB 241 AGCAACCTGGGCTGCTGGGCACTCTCGCGGTGCTGGCCCGCAGACTGCTGCGG 300
 QY 424 CACCTGGCGGCAAGGCGCGCGGCAAGTCTCCGAAGCTATGCTATGGCACTCC 483
 DB 301 CACCTGGCGGCAAGGCGCGCGGCAAGTCTCCGAAGCTATGCTATGGCACTCC 360
 QY 484 AGCTCTTCAAGAGCAGAGAGGTAGCTGCGCTGCGCGGAGTCAACAGTTCGCA 543
 DB 361 AGCTCTTCAAGAGCAGAGAGGTAGCTGCGCTGCGCGGAGTCAACAGTTCGCA 420
 QY 544 AATTCTCAGAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGCGCGAGT 603
 DB 421 AATTCTCAGAGGCTCAGTGGAGACAGGCTCCCCCAACCAAGCGGCGCGAGT 480
 QY 604 CGGGGCGGCGCGAGTGTGTCAGACGCGCGCGGAGAGAGGCGCGCAGCGGAG 663
 DB 481 CGGGGCGGCGCGAGTGTGTCAGACGCGCGCGGAGAGAGGCGCGCAGCGGAG 540
 QY 664 CAGCAGTCAAGAGCGCGCAGACCTTCTCTGAAAGGCAAAATGACTGTGATCCGGCTC 723
 DB 541 CAGCAGTCAAGAGCGCGCAGACCTTCTCTGAAAGGCAAAATGACTGTGATCCGGCTC 600
 QY 724 CGGGTTCAGACAGATGCTCGAGCATGGGCGCAGCTTGGAGCAAGGCGGTGCGG 783
 DB 601 CGGGTTCAGACAGATGCTCGAGCATGGGCGCAGCTTGGAGCAAGGCGGTGCGG 660
 QY 784 CGGGCCCGCAGGCGTGGCGCGGCGAGCTGAGAGTGTGGGAGGCGCACCGAGTGGCGG 843
 DB 661 CGGGCCCGCAGGCGTGGCGCGGCGAGCTGAGAGTGTGGGAGGCGCACCGAGTGGCGG 720
 QY 844 TCAAGGAGACTGGGCTCTGTGTGTGATCATCAAGTCTCTCAAGTCTCTATCTGAG 903

DB 721 TCAAGGAGACTGGGCTCTGTGTGTGATCATCAAGTCTCTCAAGTCTCTATCTGAG 780
 QY 904 GCCTTCGGGGCGCAGTACCTGAGTGGCGGCGCTGTCAGAGGCGCTGGGGCGGTTCCTG 963
 DB 781 GCCTTCGGGGCGCAGTACCTGAGTGGCGGCGCTGTCAGAGGCGCTGGGGCGGTTCCTG 840
 QY 964 ACTGAGGCGCTGCGAGAGCTGTGGCGCGGAGGCTGTTCGCTGCTGTCACTGTGAT 1023
 DB 841 ACTGAGGCGCTGCGAGAGCTGTGGCGCGGAGGCTGTTCGCTGCTGTCACTGTGAT 900
 QY 1024 GAGGCTGACTATGAGCTGCGCGCGCGCTTCTGATGAGAGAGAGAGGCGGCGG 1083
 DB 901 GAGGCTGACTATGAGCTGCGCGCGCGCTTCTGATGAGAGAGAGAGGCGGCGG 960
 QY 1084 CGCCCGACAGAGGCTCTGCA 1104
 DB 961 CGCCCGACAGAGGCTCTGCA 981

RESULT 10
 AAF27407
 ID AAF27407 standard; cDNA; 909 BP.

XX AAF27407;
 AC 24-APR-2001 (first entry)
 XX
 DE Human apoptosis-associated factor NT2RM1000558 partial cDNA, SEQ ID NO:1.
 XX
 KW Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
 KM DED; caspase family cleavage domain; pro-apoptotic; drug screening;
 KM cell proliferation; ischemic disease; chronic viral disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200104300-A1.
 XX
 PD 18-JAN-2001.
 XX
 PF 06-JUL-2000; 2000MO-JP004516.
 XX
 PR 08-JUL-1999; 99JP-00194179.
 PR 18-OCT-1999; 99US-0159586P.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
 XX WPI; 2001-138348/14.
 DR P-PSDB; AAB60386.
 XX
 PT Polynucleotide encoding an apoptosis-associated factor protein with death
 PT effector domain and caspase family-cleavage domain, useful in regulating
 PT diseases with cell proliferation.
 XX
 PS Claim 1; Page 41-43; 53pp; Japanese.
 XX
 CC The invention relates to a novel human apoptosis-associated factor
 CC (AAB60386, AAB60387), designated NT2RM1000558, which contains a death
 CC effector domain (DED) and a caspase family cleavage domain and is capable
 CC of inducing apoptosis in cells. The invention also relates to nucleic
 CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
 CC (particularly dominant negative variants); vectors and host cells
 CC comprising a nucleic acid which encodes an apoptosis-associated factor
 CC of the invention; the recombinant production of the protein; an antibody
 CC against the protein; and methods of screening for compounds which can
 CC regulate apoptosis. The apoptosis-related factor is useful in regulating
 CC diseases associated with cell proliferation and in screening drug
 CC candidates e.g., for regulating cell proliferation or cell death in
 CC ischemic diseases and chronic viral diseases. The present sequence
 CC represents cDNA encoding a substantial proportion of the human apoptosis-
 CC associated factor NT2RM1000558


```
DB 141 CCCCTTGACAGCCCCCTCCACAGAGATGTGGGCTCTGAGGCTTAACCATTTCCAGCTGAG 200
QY 1237 TTTCCTTCCAGACTCCTCTCTACCCCCAGGATGTGCCCCCTTAGCCTCCGAGAGGGGGGCG 1296
DB 201 TTTCCTTCCAGACTCCTCTCTACCCCCAGGATGTGCCCCCTTAGCCTCCGAGAGGGGGGCG 260
QY 1297 TGGGCGCTGTATCTCAGAAAGGAGGGGACAGCTACACACTCACCAAGAGCCCCCTTGAC 1356
DB 261 TGGGCGCTGTATCTCAGAAAGGAGGGGACAGCTACACACTCACCAAGAGCCCCCTTGAC 320
QY 1357 ATTGTATCTGTATCTTTGGGCTGTCTGCACTGTGACAGGTGACACACTGCTGCTCATGCTC 1416
DB 321 ATTGTATCTGTATCTTTGGGCTGTCTGCACTGTGACAGGTGACACACTGCTGCTCATGCTC 380
QY 1417 ACACGTGCCCCGTGAGATCTTCCCTGGGCTCTGCCCCCTGGCTTCCAGAGACAC 1476
DB 381 ACACGTGCCCCGTGAGATCTTCCCTGGGCTCTGCCCCCTGGCTTCCAGAGACAC 440
QY 1477 TTCTTTGGCCTTAAGGGCTTCTCTCTCAGACCTTAAATTGACCAACAACCACTGGGCT 1536
DB 441 TTCTTTGGCCTTAAGGGCTTCTCTCTCAGACCTTAAATTGACCAACAACCACTGGGCT 500
QY 1537 TCAGCCACATCAGTGGGCACTGAGAGCTGGGGTGACATGGGGGCTGCTCACTTTGCCAC 1596
DB 501 TCAGCCACATCAGTGGGCACTGAGAGCTGGGGTGACATGGGGGCTGCTCACTTTGCCAC 560
QY 1597 ACATCTCCAGCCAGCCAGGGGCTGCCCAGCACTTCAATTTCACACCTGACCTTCCTAC 1656
DB 561 ACATCTCCAGCCAGCCAGGGGCTGCCCAGCACTTCAATTTCACACCTGACCTTCCTAC 620
QY 1657 TTCCCCCTGCTGTCCAGAGCTGAACAATAGACTTGACCTTGGATGTGCACTGGAGTGTCA 1716
DB 621 TTCCCCCTGCTGTCCAGAGCTGAACAATAGACTTGACCTTGGATGTGCACTGGAGTGTCA 680
QY 1717 CATGGAGTGTATGCGAGCATCATACCAAGGCTTACTGTTGCAATGAGGAGCAAAACCA 1776
DB 681 CATGGAGTGTATGCGAGCATCATACCAAGGCTTACTGTTGCAATGAGGAGCAAAACCA 740
QY 1777 GTAAGACGCACTTCTTGGAAGGGGAATGCAAAAGGCTTTGGGGGTGATGGAAGAACCT 1836
DB 741 GTAAGACGCACTTCTTGGAAGGGGAATGCAAAAGGCTTTGGGGGTGATGGAAGAACCT 800
QY 1837 TTATCAATGATATCAATTAATGCTGCGCTGGAAAGGGCATAGGTGGG 1883
DB 801 TTATCAATGATATCAATTAATGCTGCGCTGGAAAGGGCATAGGTGGG 847

RESULT 12
AAS62603
ID AAS62603 standard; cDNA; 1966 BP.
XX AAS62603;
XX
AC 14-FEB-2002 (first entry)
XX
DT 14-FEB-2002 (first entry)
XX
DE cDNA sequence #390 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
KW immunosuppressive; antineumatic; ss.
XX
OS Homo sapiens.
XX
PN WO200177291-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010485.
XX
PR 06-APR-2000; 2000US-0195604P.
XX
```

```
PA (GEMV ) GENETICS INST INC.
XX
PI Mong GG, Clark HF, Reichel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulnora K, Graham JR;
XX
DR WPI; 2002-010900/01.
XX
PT New polynucleotides encoding secreted proteins useful for treating e.g.
PT asthma, HIV and Crohn's disease.
XX
PS Claim 1; Page 280; 391pp; English.
XX
CC The present invention relates to the isolation of novel cDNA sequences
CC which encode human secreted proteins. The cDNA sequences have been
CC derived from a variety of human tissues. The invention also provides a
CC method for producing proteins from these polynucleotide sequences. The
CC proteins are useful for identifying compounds that modulate their
CC activity and production, and the cell is also useful for identifying
CC compounds that modulate expression of the polynucleotide sequences
CC encoding the secreted proteins. The sequences of the invention are useful
CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
CC immune deficiency disorders (e.g. severe combined immunodeficiency
CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis) and
CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
CC the invention are also useful in gene therapy. AAS62214-AAS62838
CC represent the cDNA sequences of the invention that encode for novel human
CC secreted proteins
XX
SQ Sequence 1966 BP; 356 A; 618 C; 614 G; 378 T; 0 U; 0 Other;
```

```
Query Match 37.5%; Score 707; DB 6; Length 1966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1177 CCCCTTGACAGCCCCCTCCACAGAGATGTGGGCTCTGAGGCTTAACCATTTCCAGCTGAG 1236
DB 1260 CCCCTTGACAGCCCCCTCCACAGAGATGTGGGCTCTGAGGCTTAACCATTTCCAGCTGAG 1319
QY 1237 TTTCCTTCCAGACTCTCTCTACCCCCAGGATGTGCCCCCTTACCTCCGAGGCGGGGCG 1296
DB 1320 TTTCCTTCCAGACTCTCTCTACCCCCAGGATGTGCCCCCTTACCTCCGAGGCGGGGCG 1379
QY 1297 TGGGCGCTGTATCTCAGAAAGGAGGGGACAGCTACACACTCACCAAGAGCCCCCTTGAC 1356
DB 1380 TGGGCGCTGTATCTCAGAAAGGAGGGGACAGCTACACACTCACCAAGAGCCCCCTTGAC 1439
QY 1357 ATTGTATCTGTATCTTTGGGCTGTCTGCACTGTGACAGGTGACACACTGCTCATGCTC 1416
DB 1440 ATTGTATCTGTATCTTTGGGCTGTCTGCACTGTGACAGGTGACACACTGCTCATGCTC 1499
QY 1417 ACACGTGCCCCGTGAGATCTTCCCTGGGCTGTGCCCCCTGCTTCCAGACACAC 1476
DB 1500 ACACGTGCCCCGTGAGATCTTCCCTGGGCTGTGCCCCCTGCTTCCAGACACAC 1559
QY 1477 TTCTTTGGCCTTAAGGGCTTCTCTCAGACCTTAAATTGACCAACAACCACTGGGCT 1536
DB 1560 TTCTTTGGCCTTAAGGGCTTCTCTCAGACCTTAAATTGACCAACAACCACTGGGCT 1619
QY 1537 TCAGCCACATCAGTGGGCACTGAGAGCTGGGGTGACATGGGGGCTGCTCACTTTGCCAC 1596
DB 1620 TCAGCCACATCAGTGGGCACTGAGAGCTGGGGTGACATGGGGGCTGCTCACTTTGCCAC 1679
QY 1597 ACATCTCCAGCCAGCCAGGGGCTGCCCAGCACTTCAATTTCAGACACTGACTCTCTAC 1656
DB 1680 ACATCTCCAGCCAGCCAGGGGCTGCCCAGCACTTCAATTTCAGACACTGACTCTCTAC 1739
QY 1657 TTCCCCCTGCTGTCCAGAGCTGAACATAGACTTGACCTTGGATGTCACTGGAGTGTCA 1716
DB 1740 TTCCCCCTGCTGTCCAGAGCTGAACATAGACTTGACCTTGGATGTCACTGGAGTGTCA 1799
QY 1717 CATGGAGTGTATGCGAGCATCATACCAAGGCTTACTGTTGCAATGAGGAGCAAAACCA 1776
```


DB 1800 CATTGGAGTGTATTGACAGCATCATCAAGAGCCCTTACTGTTGACATCTGGGGCCAAAACCA 1859
QY 1777 GTAAACAGCCACCTTCTTGGAAAGGATGCAAGAGGCTTGGGGGTGATGGAAGACCT 1836
DB 1860 GTAAACAGCCACCTTCTTGGAAAGGATGCAAGAGGCTTGGGGGTGATGGAAGACCT 1919
QY 1837 TTTTCAATGATTCATTAATTAAGTCCCTGGGAAAGGGCATAGTGGG 1883
DB 1920 TTTTCAATGATTCATTAATTAAGTCCCTGGGAAAGGGCATAGTGGG 1966

RESULT 13
AAH9646
ID AAH9646 standard; cDNA; 1570 BP.
XX
AC AAH9646;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:481.
XX
KW Human; cancer; HIV infection; human immunodeficiency virus;
KW anti-inflammation; anti-rheumatic; anti-arthritic; immunosuppressive;
KW antibacterial; endocrine; cardiac; central nervous system; viricide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiallergic; haemostatic; vulnary; antilucer; osteopathic; eczema;
KW dermatological; antiallergic; antiallergic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antinflammatory; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US035017.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
XX
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-457603/49.
DR P-PSDB; AAM25705.
XX
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
XX
XX
PS Claim 1; Page 552; 1217pp; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and cells
CC they are expressed in, such as: anti-inflammation; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
CC central nervous system; viricide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiallergic; haemostatic; vulnary;
CC antilucer; osteopathic; dermatological; antiallergic; antidiabetic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders
XX

Sequence 1570 BP; 311 A; 488 C; 466 G; 305 T; 0 U; 0 Other;
Query Match 36.1%; Score 680; DB 4; Length 1570;
Best Local Similarity 99.9%; Pred. No. 1.9e-305;
Matches 730; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 451 GTGTCTCCGAAGCGCTATGATGAGCACTCCAGCTCTTCAAGAGAGAGAGGATAGC 510
DB 76 GTGTCTCCGAAGCGCTATGATGAGCACTCCAGCTCTTCAAGAGAGAGAGGATAGC 135
QY 511 TGGCGTGGCGCTCGGAGATCAAGAGAGTTCTGCAAAATTCAGAGAGGTGAGAGACA 570
DB 136 TGGCGTGGCGCTCGGAGATCAAGAGAGTTCTGCAAAATTCAGAGAGGTGAGAGACA 195
QY 571 GGCTCCCCCAACCAAG 630
DB 196 GGCTCCCCCAACCAAG 255
QY 631 CGGCGCGAG 690
DB 256 CGGCGCGAG 315
QY 691 TGTGAAGGCAAG 750
DB 316 TGTGAAGGCAAG 375
QY 751 GGGCCAGCTTGGAG 810
DB 376 GGGCCAGCTTGGAG 435
QY 811 GACGTGTTGGAG 870
DB 436 GACGTGTTGGAG 495
QY 871 GACATCAAGTTCTCAG 930
DB 496 GACATCAAGTTCTCAG 555
QY 931 GCGCTGTCAG 990
DB 556 GCGCTGTCAG 615
QY 991 CGGAG 1050
DB 616 CGGAG 675
QY 1051 CGCTGTTGCTGATGAG 1110
DB 676 CGCTGTTGCTGATGAG 735
QY 1111 ACTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170
DB 736 ACTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
QY 1171 TCTTACCCCT 1181
DB 796 TCTTACCCCT 806

RESULT 14
ABZ70266
ID ABZ70266 standard; cDNA; 857 BP.
XX

AC AB270266;
 XX
 DT 25-APR-2003 (first entry)
 XX
 DE ATP synthetase 11.55 coding sequence.
 XX
 KM ATP synthetase 11.55; enzyme; cancer; cytosolic; HIV infection;
 XX anti-HIV; gene; ss.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT CDS 137..454
 FT /*tag= a
 FT /product= "ATP synthetase 11.55"
 XX
 PN CNI363656-A.
 XX
 PD 14-AUG-2002.
 XX
 PF 05-JAN-2001; 2001CN-00105031.
 XX
 PR 05-JAN-2001; 2001CN-00105031.
 XX
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI, 2002-751770/82.
 DR P-PSDB; ABP59171.
 XX
 PT Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.
 XX
 PS Claim 6; Page 25 (Disclosure); 32pp; Chinese.
 CC
 CC The present sequence is the coding sequence for ATP synthetase 11.55. The
 CC protein can be used for treating diseases such as cancer and HIV
 CC infection
 CC
 SQ Sequence 857 BP; 191 A; 279 C; 204 G; 183 T; 0 U; 0 Other;
 Query Match 34.8%; Score 656; DB 6; Length 857;
 Best Local Similarity 99.9%; Pred. No. 2.8e-294;
 Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1177 CCCCTTGACAGCCCTCCACAGAGATGTGGCTTGAAGCTTAACATTTCAGCTGAG 1236
 DB 113 CCCCTTGACAGCCCTCCACAGAGATGTGGCTTGAAGCTTAACATTTCAGCTGAG 172
 QY 1237 TTTCCTTCCAGACTCCTCCACAGAGATGTGGCTTGAAGCTTCCGAGAGGAGG 1296
 DB 173 TTTCCTTCCAGACTCCTCCACAGAGATGTGGCTTGAAGCTTCCGAGAGGAGG 232
 QY 1297 TGGGCTGTATCTCAGAGAGGAGGAGACAGCTACACACTACCAAGGCCCTGAC 1356
 DB 233 TGGGCTGTATCTCAGAGAGGAGGAGACAGCTACACACTACCAAGGCCCTGAC 292
 QY 1357 ATTGATCTGTGATCTTGGGCTGTCTGACATGTACAGAGTGCACACTGGCTATGCTC 1416
 DB 293 ATTGATCTGTGATCTTGGGCTGTCTGACATGTACAGAGTGCACACTGGCTATGCTC 352
 QY 1417 ACACTGCCCCCTGAGATCTTCCCTGGGCTTGCCTGGCCCTGCTTCCAGACACAC 1476
 DB 353 ACACTGCCCCCTGAGATCTTCCCTGGGCTTGCCTGGCCCTGCTTCCAGACACAC 412
 QY 1477 TTCTTGGCTTAAGGGCTTCTCTCTCAGACCTTAAATTGACCAACCAACTGGGCT 1536
 DB 413 TTCTTGGCTTAAGGGCTTCTCTCTCAGACCTTAAATTGACCAACCAACTGGGCT 472
 QY 1537 TCAGCCCATCATGAGGAGCTGAGAGCTGGGGTGCACATGGGGCTGCTACCTTGCCAC 1596
 DB 473 TCAGCCCATCATGAGGAGCTGAGAGCTGGGGTGCACATGGGGCTGCTACCTTGCCAC 532

QY 1597 ACATCTCCAGCCAGCAGAGGCGCTGCCACGCTTCAATTACAGACCTGACTCTCCACC 1656
 DB 533 ACATCTCCAGCCAGCAGAGGCGCTGCCACGCTTCAATTACAGACCTGACTCTCCACC 592
 QY 1657 TTCCCCCTGCTGCTCCAGAGCTGAACATAGACTTTGACCTTGATGTCACTGGAGTCA 1716
 DB 593 TTCCCCCTGCTGCTCCAGAGCTGAACATAGACTTTGACCTTGATGTCACTGGAGTCA 652
 QY 1717 CATGGAGTGTATGGAGCATCATACCAAGGCTACTGTGTGACATGGGGCCCAACCA 1776
 DB 653 CATGGAGTGTATGGAGCATCATACCAAGGCTACTGTGTGACATGGGGCCCAACCA 712
 QY 1777 GTAACAGCCACCTTCTTGGAAAAGGATGCAAGGCTTTGGGGGTGATGAAAAGACT 1836
 DB 713 GTAACAGCCACCTTCTTGGAAAAGGATGCAAGGCTTTGGGGGTGATGAAAAGACT 772
 QY 1837 TTTACAAATGATACCAATTAACTGCCCTTGGAAAAGGCAATAGTGGG 1883
 DB 773 TTTACAAATGATACCAATTAACTGCCCTTGGAAAAGGCAATAGTGGG 819

RESULT 15

AAH07925 standard; cDNA; 626 BP.

AAH07925;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:4760.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

26-JUL-2000; 2000EP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

11-JAN-2000; 2000JP-00118776.

02-MAY-2000; 2000JP-00183767.

09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Claim 1; SEQ ID NO 4760; 2537bp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification, where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the

CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

SO Sequence 626 BP; 102 A; 207 C; 216 G; 98 T; 0 U; 3 Other;

Query Match 22.4%; Score 421; DB 4; Length 626;
Best Local Similarity 99.8%; Pred. No. 4.8e-185;

Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 37 AACAAATGGCTGGCGGCGCTGCGCCAGAGTCAATGAGCCGCAAAATCTGGCCGGTTCTG 96
Db 1 AACAAATGGCTGGCGGCGCTGCGCCAGAGTCAATGAGCCGCAAAATCTGGCGGTTCTG 60
QY 97 AGCTTGTCCGCTCCCTCCCTCCCGGGAATGCGCTATCCGGGTGCAACCCCGCCCGTGC 156
Db 61 AGCTTGTCCGCTCCCTCCCTCCCGGGAATGCGCTATCCGGGTGCAACCCCGCCCGTGC 120
QY 157 TGGGAGAGAGATGATGCTCTGAGCTAATACTAGCGGATGCTGCTTCAACCGTATGTTGAG 216
Db 121 TGGGAGAGAGATGATGCTCTGAGCTAATACTAGCGGATGCTGCTTCAACCGTATGTTGAG 180
QY 217 GTGGTGGGGGGGGAATGAGTCCGAGTCCGAGTCCCTGGCCCTTTCGTCTGATGAG 276
Db 181 GTGGTGGGGGGGGAATGAGTCCGAGTCCGAGTCCCTGGCCCTTTCGTCTGATGAG 240
QY 277 GCTCCTGCGCGCCGCGAGAGCTTAAAGCCCGGAGCGGCTTAAAGCTCTGCTGAG 316
Db 241 GCTCCTGCGCGCCGCGAGAGCTTAAAGCCCGGAGCGGCTTAAAGCTCTGCTGAG 300
QY 337 CTGGAGCGCGCGCGGAGTGCAGTGCAGAGCAACTGCGGTGCTGAGGCACTCTGCGC 396
Db 301 CTGGAGCGCGCGCGGAGTGCAGTGCAGAGCAACTGCGGTGCTGAGGCAACTCTGCGC 360
QY 397 GTGCTGGCGCGCGAGCACTGCTGCGGCACTGCGGCGGAGGCGGCGCGAGTGTCT 456
Db 361 GTGCTGGCGCGCGAGCACTGCTGCGGCACTGCGGCGGAGGCGGCGCGAGTGTCT 420
QY 457 CCGAAGCGCTATAGTATGCACTTCCTTTCAAGAGGAGGAGGTA 508
Db 421 CCGAAGCGCTATAGTATGCACTTCCTTTCAAGAGGAGGAGGTA 472

RESULT 16
AA160747/c
ID AA160747 standard; cDNA; 1042 BP.

XX AA160747;

XX 22-OCT-2001 (first entry)

DE Human polynucleotide SEQ ID NO 4736.

XX Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukemia; ss.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000MO-US034263.
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
PI Zhou P, Goodrich R, Drmanac RT;

DR WPI; 2001-442253/47.
XX P-PSDB; AAM41591.

PS Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.

PS Claim 1, SEQ ID NO 4736; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with neurotrophic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression.
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

SO Sequence 1042 BP; 182 A; 318 C; 334 G; 208 T; 0 U; 0 Other;

Query Match 19.5%; Score 368; DB 4; Length 1042;

Best Local Similarity 99.4%; Pred. No. 2.1e-160;

Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 37 AACAAATGGCTGGCGGCGCTGCGCCAGAGTCAATGAGCCGCAAAATCTGGCCGGTTCTG 96
Db 1042 AACAAATGGCTGGCGGCGCTGCGCCAGAGTCAATGAGCCGCAAAATCTGGCCGGTTCTG 983
QY 97 AGCTTGTCCGCTCCCTCCCTCCCGGGAATGCGCTATCCGGGTGCAACCCCGCCCGTGC 156
Db 982 AGCTTGTCCGCTCCCTCCCTCCCGGGAATGCGCTATCCGGGTGCAACCCCGCCCGTGC 923
QY 157 TGGGAGAGAGATGATGCTCTGAGCTAATACTAGCGGATGCTGCTTCAACCGTATGTTGAG 216
Db 922 TGGGAGAGAGATGATGCTCTGAGCTAATACTAGCGGATGCTGCTTCAACCGTATGTTGAG 863
QY 217 GTGGTGGGGGGGGAATGAGTCCGAGTCCGAGTCCCTGGCCCTTTCGTCTGATGAG 276
Db 862 GTGGTGGGGGGGGAATGAGTCCGAGTCCGAGTCCCTGGCCCTTTCGTCTGATGAG 803
QY 277 GCTCCTGCGCGCGCGGAGGCTTAAAGCCCGGAGCGGCTTAAAGCTCTGCTGAG 316
Db 802 GCTCCTGCGCGCGCGGAGGCTTAAAGCCCGGAGCGGCTTAAAGCTCTGCTGAG 743
QY 337 CTGGAGCGCGCGGAGTGCAGTGCAGAGCAACTGCGGTGCTGAGGCACTCTGCGC 396
Db 742 CTGGAGCGCGCGGAGTGCAGTGCAGAGCAACTGCGGTGCTGAGGCAACTCTGCGC 683
QY 397 GTGCTGGCGCGCGAGCACTGCTGCGGCACTGCGGCGGAGGCGGCGCGAGTGTCT 456

```

|||||
Db      682 GTGTGGCCCGCCAGACCTGCTGCGCACCCTGGCGGCAAGCCGCCAGTGTCT 623
Qy      457 CCAAGACGCTATAGCTATGGACCTCCAGCTCTTCAAGAAGAGAGAGGGTATGCTGCCGT 516
Db      622 CCAAGACGCTATAGCTATGGACCTCCAGCTCTTCAAGAAGAGAGAGGGTATGCTGCCGT 563
Qy      517 CGCGGTGGCAGTCAAGCAGTTCTGCAAAATTCAGCAGGG 557
Db      562 CGCGGTGGCAGTCAAGCAGTTCTGCAAAATTCAGCAGGG 522

RESULT 17
AA158961 standard; cDNA; 1067 BP.
AA158961;
22-OCT-2001 (first entry)
Human polynucleotide SEQ ID NO 1164.
Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukemia; ss.
Homo sapiens.
MO200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US034263.
23-DEC-1999; 99US-00471275.
21-JAN-2000; 2000US-00488725.
25-APR-2000; 2000US-00552317.
20-JUN-2000; 2000US-00598042.
19-JUL-2000; 2000US-00620312.
03-AUG-2000; 2000US-00653450.
14-SEP-2000; 2000US-00662191.
19-OCT-2000; 2000US-00693036.
29-NOV-2000; 2000US-00727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA,
Zhou P, Goodrich R, Drmanac RT,
WPI; 2001-442253/47.
P-PSDB; AAM39805.
Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
Claim 1; SEQ ID NO 1164; 10078bp; English.
The invention relates to human nucleic acids (AA157798-AA161369) and the
encoded polypeptides (AAM38642-AAM42213) with nocotropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,

```

```

CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;
Query Match 19.5%; Score 368; DB 4; Length 1067;
Best Local Similarity 99.4%; Pred. No. 2.1e-160;
Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 37 AACAAATGGCTGGCGCCGTGCGCCAGAGTCAATGCAAGCAGATCTGCGCGGTTCTG 96
Db 64 AACAAATGGCTGGCGCCGTGCGCCAGAGTCAATGCAAGCAGATCTGCGCGGTTCTG 123
Qy 97 AGCTGTTCGGCTCCCTCCCGGGAAATGGCGTATCCGGGTGACCCCGGCGCCGTG 156
Db 124 AGCTGTTCGGCTCCCTCCCGGGAAATGGCGTATCCGGGTGACCCCGGCGCCGTG 183
Qy 157 TGGAGAGAGATGATGATGCTGACTACTAAGGATGCTGCTTCAACCGATGTTGAG 216
Db 184 TGGAGAGAGATGATGATGCTGACTACTAAGGATGCTGCTTCAACCGATGTTGAG 243
Qy 217 GTGTGGCGCGGCACTGACCGAGTGCAGCTGAGCTTCTGCGCTTTCTGCTGATGAG 276
Db 244 GTGTGGCGCGGCACTGACCGAGTGCAGCTGAGCTTCTGCGCTTTCTGCTGATGAG 303
Qy 277 GTCTCTGGCGCCCGCGAGGCTTAAGCCCGGGCCCGCAGCGGCTTAAGCTCTGCTGAG 336
Db 304 GTCTCTGGCGCCCGCGAGGCTTAATCCGGGCCCGCAGCGGCTTAAGCTCTGCTGAG 363
Qy 337 CTGAGCGCGCGGCGAGTGCAGTGCAGAGAGCACTGCGCTGCTGAGGCAATCTCTGCG 396
Db 364 CTGAGCGCGCGGCGAGTGCAGTGCAGAGAGCACTGCGCTGCTGAGGCAATCTCTGCG 423
Qy 397 GTGTGGCGCGGCACTGCTGCTGCGCACCTGCGCGCAAGCGCGCGCGCAATGCT 456
Db 424 GTGTGGCGCGGCACTGCTGCTGCGCACCTGCGCGCAAGCGCGCGCGCAATGCT 483
Qy 457 CCAAGACGCTATAGCTATGGACCTCCAGCTCTTCAAGAAGAGAGAGGGTATGCTGCCGT 516
Db 484 CCAAGACGCTATAGCTATGGACCTCCAGCTCTTCAAGAAGAGAGAGGGTATGCTGCCGT 543
Qy 517 CGCGGTGGCAGTCAAGCAGTTCTGCAAAATTCAGCAGGG 557
Db 544 CGCGGTGGCAGTCAAGCAGTTCTGCAAAATTCAGCAGGG 584

RESULT 18
AD099183
ID AD099183 standard; cDNA; 1067 BP.
AC AD099183;
23-SEP-2004 (first entry)
DNA encoding human GPCR-like protein seqid 853.
Kw ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
Kw antidiabetic; GPCR-like protein; ophthalmic disorder;
Kw neurological disorder; immunological disorder; nephritic disorder;
Kw hormonal dysfunction; cancer; atherosclerosis; diabetes;
Kw molecular weight marker; food supplement; human; ss.
OS Homo sapiens.
PN US6569662-B1.
XX US6569662-B1.
XX 27-MAY-2003.
PF 19-JUL-2000; 2000US-00620312.
XX 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.

```

```
XX (HYSE-) HYSEQ INC.
PA Tang YT, Zhou P, Drmanac RT;
XX WPI; 2001-442255/47.
XX
XX New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT useful for treating diseases of ophthalmic, neurological, immunological
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
XX and diabetes.
XX
XX Example 2; SEQ ID NO 853; 92pp; English.
XX
XX The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (SI) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translated or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.
XX
SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;

Query Match 19.5%; Score 368; DB 5; Length 1067;
Best Local Similarity 99.4%; Pred. No. 2.1e-160;
Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 37 AACAAAGTGGCTGGCGCCGTCGAGGATCGATCGAGCCAGAAATCTGCGCGGTTCTG 96
DB 64 AACAAAGTGGCTGGCGCCGTCGAGGATCGATCGAGCCAGAAATCTGCGCGGTTCTG 123
QY 97 AGCTTGTTCGCTCCCTCCCTCCCGGGAATGCGCTATCCGCGGTGACCCCGGCCGTGC 156
DB 124 AGCTTGTTCGCTCCCTCCCTCCCGGGAATGCGCTATCCGCGGTGACCCCGGCCGTGC 183
QY 157 TGGGAGAGAGATAGTACCTGAGCTACTAGGAGATGCTGTCGCTTACCGGTATGTTGAG 216
DB 184 TGGGAGAGAGATAGTACCTGAGCTACTAGGAGATGCTGTCGCTTACCGGTATGTTGAG 243
QY 217 GTGCTGGCGCGGAGCACTGAGCTGAGCTGAGCTCTGCTGATGAG 276
DB 244 GTGCTGGCGCGGAGCACTGAGCTGAGCTGAGCTCTGCTGATGAG 303
QY 277 GCTCTGGCGCGCGGAGCTTATGCCCGGAGCCCGGAGCGGCTTATGCTGCTGAGAG 336
DB 304 GCTCTGGCGCGCGGAGCTTATGCCCGGAGCCCGGAGCGGCTTATGCTGCTGAGAG 363
QY 337 CTGAGAGCGCGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 396
DB 364 CTGAGAGCGCGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 423
QY 397 GTGCTGGCGCGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 456
DB 424 GTGCTGGCGCGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 483
QY 457 CCAAGACGCTATAGCTATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 516
DB 484 CCAAGACGCTATAGCTATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 543
QY 517 CGCGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 557
DB 544 CGCGCTGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 584

RESULT 19
ADB48943
ID ADB48943 standard; cDNA; 1067 BP.
```

```
XX ADB48943;
AC 04-DEC-2003 (first entry)
XX
XX Novel human cDNA SEQ ID NO 853.
XX
XX ss; cancer; neurodegenerative disease; human.
XX Homo sapiens.
XX
XX US2003104529-A1.
XX
XX 05-JUN-2003.
XX
XX 04-JAN-2002; 2002US-00037270.
XX
XX 21-JUN-2000; 2000US-00488725.
XX 25-APR-2000; 2000US-00552317.
XX 19-JUL-2000; 2000US-00620312.
XX
XX (ZHOU/) ZHOU P.
XX (TANG/) TANG Y T.
XX (LIUC/) LIU C.
XX (ASUN/) ASUNDI V.
XX (DRMA/) DRMANAC R T.
XX
XX Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
XX WPI; 2003-678194/64.
XX
XX New polynucleotide, useful for treating diseases e.g., cancer or
XX neurodegenerative diseases.
XX
XX Claim 1; SEQ ID NO 853; 99pp; English.
XX
XX The invention relates to a polynucleotide comprising a sequence given in
XX the specification, or its mature protein-coding portion, or its
XX complement. The polynucleotide is useful for treating diseases e.g.,
XX cancer or neurodegenerative diseases and many others listed in the
XX specification. The present sequence represents a novel human cDNA. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?docid=2003104529.
XX
SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;

Query Match 19.5%; Score 368; DB 9; Length 1067;
Best Local Similarity 99.4%; Pred. No. 2.1e-160;
Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 37 AACAAAGTGGCTGGCGCCGTCGAGGATCGATCGAGCCAGAAATCTGCGCGGTTCTG 96
DB 64 AACAAAGTGGCTGGCGCCGTCGAGGATCGATCGAGCCAGAAATCTGCGCGGTTCTG 123
QY 97 AGCTTGTTCGCTCCCTCCCTCCCGGGAATGCGCTATCCGCGGTGACCCCGGCCGTGC 156
DB 124 AGCTTGTTCGCTCCCTCCCTCCCGGGAATGCGCTATCCGCGGTGACCCCGGCCGTGC 183
QY 157 TGGGAGAGAGATAGTACCTGAGCTACTAGGAGATGCTGTCGCTTACCGGTATGTTGAG 216
DB 184 TGGGAGAGAGATAGTACCTGAGCTACTAGGAGATGCTGTCGCTTACCGGTATGTTGAG 243
QY 217 GTGCTGGCGCGGAGCACTGAGCTGAGCTGAGCTGAGCTCTGCTGATGAG 276
DB 244 GTGCTGGCGCGGAGCACTGAGCTGAGCTGAGCTGAGCTCTGCTGATGAG 303
QY 277 GCTCTGGCGCGCGGAGCTTATGCCCGGAGCCCGGAGCGGCTTATGCTGCTGAGAG 336
DB 304 GCTCTGGCGCGCGGAGCTTATGCCCGGAGCCCGGAGCGGCTTATGCTGCTGAGAG 363
QY 337 CTGAGAGCGCGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 396
```

Db 364 CTGAGAGCCGCGGAGTGCAGAGAGCAACCTGCGCTGCTGGGCAACTCTCCGCG 423
Qy 397 GTGCTGGCCCGCCAGACAGTCTGCTGCGACCTGGCGGCAAGCGCGCCGAGTGTCT 456
Db 424 GTGCTGGCCCGCCAGACAGTCTGCTGCGACCTGGCGGCAAGCGCGCCGAGTGTCT 483
Qy 457 CGAAGAGCTATAGCTATAGGCACTTCCAGCTTTCAAGAGAGAGAGGAGTGGCTGCT 516
Db 484 CCAAGAGCTATAGCTATAGGCACTTCCAGCTTTCAAGAGAGAGAGGAGTGGCTGCT 543
Qy 517 CGCGCTGGAGCTACAGCACTTCTGCAAAATTTCTAGAGAGG 557
Db 544 CGCGCTGGAGCTACAGCACTTCTGCAAAATTTCTAGAGAGG 584

RESULT 20
ADQ18808/c
ID ADQ18808 standard; DNA; 111084 BP.
XX
AC ADQ18808;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 1627.
XX
KM soft tissue sarcoma; cytoskeletal; gene therapy; vaccine; screening; human;
XX da.
XX Homo sapiens.
XX MO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003MO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
PS Example 2; SEQ ID NO 1627; 210bp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytoskeletal applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
PS Sequence 111084 BP; 24747 A; 31184 C; 30978 G; 24175 T; 0 U; 0 Other;
SQ

Query Match 16.8%; Score 316; DB 12; Length 111084;
Best Local Similarity 99.7%; Pred. No. 2,6e-116;
Matches 366; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 85 GGCGGGTTTGAGCTTGTTCGCGCTCCCTCCCGGGAATGGCGCTATCGGGTGCACC 144
|||||

Db 8901 GGCGGGTTTGAGCTTGTTCGCGCTCCCTCCCGGGAATGGCGCTATCGGGTGCACC 8842
Qy 145 CGCGCCCGGCTGGAGAGAGAGATGATGCTGCTGACTACTACGGAGATGCTGTGCCTTAC 204
Db 8841 CGCGCCCGGCTGGAGAGAGATGATGCTGCTGACTACTACGGAGATGCTGTGCCTTAC 8782
Qy 205 CGATATGTCAGAGTGTGTGGCGGCACTGACCGAGTGCAGTGTGAGCTCTGGCCCTTT 264
Db 8781 CGATATGTCAGAGTGTGTGGCGGCACTGACCGAGTGCAGTGTGAGCTCTGGCCCTTT 8722
Qy 265 CTGCTGATGAGCTCTCTGCGCGCCCGGAGGCTTACCTGGCGCGGCGCTTACAG 324
Db 8721 CTGCTGATGAGCTCTCTGCGCGCCCGGAGGCTTACCTGGCGCGGCGCTTACAG 8662
Qy 325 CTCTGCTGAGCTGAGAGCGCGCGGAGTGTGGGAGAGCAACTGGGGCTGTGGGG 384
Db 8661 CTCTGCTGAGCTGAGAGCGCGCGGAGTGTGGGAGAGCAACTGGGGCTGTGGGG 8602
Qy 385 CAATCTCTGCGCGTGTGCTGCGCCGCAAGACTGTGCTGCGCGCAAGCGCGC 444
Db 8601 CAATCTCTGCGCGTGTGCTGCGCCGCAAGACTGTGCTGCGCGCAAGCGCGC 8542
Qy 445 CGGCGAG 451
Db 8541 CGGCGAG 8535

RESULT 21
ACH41209
ID ACH41209 standard; cDNA; 451 BP.
XX
AC ACH41209;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human foetal brain cDNA #2576.
XX
KM Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KM genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-Apr-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 28421; 44bp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH412789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for

XX (HELI-) HELIX RES INST.
PA Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
PI WPI, 2001-093989/11.
DR
XX
PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
XX gene therapy or as candidate target molecules in drug development.
PS Claim 4; SEQ ID NO 371; 609pp + Sequence Listing; English.
XX
XX This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC which encode human secretory or membrane proteins represented by AAB88317
CC - AAB88419. Included in the invention are primers AAF93917 - AAF94295 and
CC AAF62232 - AAF62235 which are used to isolate the cDNA sequences of the
CC invention. The invention also includes methods for the production of
CC antibodies directed against the proteins, and cDNA sequences, which can
CC be used in vaccines. The polynucleotide sequences can be used in gene
CC therapy. The polynucleotide sequences and the proteins they encode may be
CC used in the prevention, treatment and diagnosis of diseases associated
CC with inappropriate secretory protein/membrane protein expression. The
CC nucleic acids and complementary sequences may also be used as DNA probes
CC in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect
CC and quantitate the presence of similar nucleic acid sequences in samples.
CC They may also be used to study the expression and function of secretory
CC proteins/membrane polypeptides and their role in metabolism. The
CC polypeptides may be used as antigens in the production of antibodies
CC against them and in assays to identify modulators (agonists and
CC antagonists) of expression and activity. The antibodies and antagonists
CC may also be used as therapeutic agents to down regulate expression and
CC activity. The antibodies may also be used as diagnostic agents for
CC detecting the presence of the polypeptides in samples (e.g. by enzyme
CC linked immunosorbent assay (ELISA)). Examples of diseases which may be
CC treated include rheumatoid arthritis and diabetes
XX
SQ Sequence 603 BP; 114 A; 189 C; 187 G; 103 T; 0 U; 10 Other;
Query Match 14.2%; Score 268; DB 5; Length 603;
Best Local Similarity 100.0%; Pred. No. 6.3e-114;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 CCAAAACAATGAGCTGGCGCGCTGCGCCAGAGATCATCGACGCAAGATCTGGCGCGGT 92
DB 32 CCAAAACAATGAGCTGGCGCGCTGCGCCAGAGATCATCGACGCAAGATCTGGCGCGGT 91
QY 93 TCTGAGCTTGTTCGCTCCCTCCCTCCCGGGAATGCGCTATCCGGGTGACCCCGGCCCC 152
DB 92 TCTGAGCTTGTTCGCTCCCTCCCTCCCGGGAATGCGCTATCCGGGTGACCCCGGCCCC 151
QY 153 GTGCTGGAGAGAGATGAGTGCCTGACTACGGGATCGTGCCTTACCGTATGTT 212
DB 152 GTGCTGGAGAGAGATGAGTGCCTGACTACGGGATCGTGCCTTACCGTATGTT 211
QY 213 CGAGTGTGTGGGGGGGCACTGACCGAGTGCAGAGCTGAGAGCTTCTGCTGGA 272
DB 212 CGAGTGTGTGGGGGGGCACTGACCGAGTGCAGAGCTGAGAGCTTCTGCTGGA 271
QY 273 TGAGGCTCTGGCGCGCGCCGAGGCTTAA 300
DB 272 TGAGGCTCTGGCGCGCGCCGAGGCTTAA 299
RESULT 24
AAD40075
ID AAD40075 standard; cDNA; 303 BP.
XX
XX AAD40075;
XX
XX 22-OCT-2002 (first entry)
XX
XX Human DED4 DED (death effector domain) cDNA.

KM Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KM NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KM inflammation; allergy; autoimmunity; allograft rejection; cell division;
KM immune-based pathology; fibrosis; arthritis; graft versus host disease;
KM immunosuppressive; gene therapy; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key 1..303 Location/Qualifiers
FT CDS
FT 1..303
FT /*tag= a
FT /product= "Human DED4 DED"
FT /note= "No start and stop codon"
FT /partial
XX
XX NO200240680-A2.
XX
XX 23-MAY-2002.
XX
XX 15-NOV-2001; 2001MO-US044844.
XX
XX 17-NOV-2000; 2000US-00715893.
XX 23-JUN-2001; 2001US-0301889P.
XX
XX (BURN-) BURHAM INST.
XX
PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
PI Stemmer-Liwen F;
XX
XX WPI, 2002-500222/53.
XX P-PSDB; AAB24855.
XX
XX New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,
PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
XX Claim 18; Page 174-175; 209pp; English.
XX
XX The invention relates to an isolated polypeptide comprising a death
CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
CC is useful for identifying a binding agent, preferably a protein or a drug
CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
CC domain from DAP3, IRAK4, CTDD (Chlamydia trachomatis DP protein), DED4 or
CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
CC detecting the association of the domain and the candidate binding agent,
CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
CC spectroscopy (MS) and FPA. The invention is useful for modulating the
CC level of a cell process such as cell proliferation, cell adhesion, cell
CC stress responses, responses to microbial infection and B cell
CC immunoglobulin class switching, in particular apoptosis within a cell.
CC Antibody specifically reactive with CTDD DD of C. trachomatis, C.
CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
CC CTDD DD protein is useful for detecting a Chlamydia infection. The
CC invention is useful for modulating the activity of oncogenic proteins,
CC for treating a pathology caused by the oncogenic proteins and for
CC treating bacterial infections by modulating the activity of bacterial
CC proteins. The protein and antibody specific for it are useful for
CC discovery of drugs that suppress infection, inflammation, allergy,
CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
CC is useful for treating immune-based pathologies, pathologies associated
CC with cell division, inflammatory diseases such as sepsis, fibrosis,
CC arthritis, graft versus host disease. The invention is used in antisense
CC therapy and gene therapy. The present sequence is human DED4 DED cDNA
XX
SQ Sequence 303 BP; 41 A; 96 C; 115 G; 51 T; 0 U; 0 Other;
Query Match 13.4%; Score 252; DB 6; Length 303;
Best Local Similarity 99.7%; Pred. No. 1.8e-106;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 157 TGGAGAGAGATGAGTGCCTGACTGACCGGATGCTGTGCTTACCGTATGTTGAG 216
|||||

```

Db      1 TGGGAGAGAGATAGTGTCTGAGCTACTACGAGATGCTGTGCTTACCGGTATGTTGAG 60
QY      217 GTGGTGGGCGGCACTGACCGAGTGGAGCTGCTGCTTCTTGTCTGATGAG 276
        |||
        |||
Db      61 GTGTGGGGGGGCACTGACCGAGTGGAGCTGAGCTCTTGCTTGTCTGATGAG 120
QY      277 GCTCTGGGCGCGCCGAGAGCTTACGCCCGGCGGCTTAAGCTTCTGCTGAG 336
        |||
        |||
Db      121 GCTCTGGGCGCGCCGAGAGCTTACGCCCGGCGGCTTAAGCTTCTGCTGAG 180
QY      337 CTGAGAGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTCTGCGG 396
        |||
        |||
Db      181 CTGAGAGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTCTGCGG 240
QY      397 GTGCTGGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTCTGCT 456
        |||
        |||
Db      241 GTGCTGGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTGTCT 300
QY      457 CCA 459
        |||
        |||
Db      301 CCA 303

RESULT 25
AADS9057 standard; cDNA; 303 BP.
XX      AADS9057;
AC      AADS9057;
XX      18-DEC-2003 (first entry)
DE      Human DED4 DED cDNA.
XX      Human; death Domain; DD; death effector domain; DED; cell proliferation;
KM      Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;
KM      neural growth factor receptor-interacting death domain; cell adhesion;
KM      vasotropic; microbial infection; inflammation; allograft rejection; CTDD;
KM      cell stress response; benign prostatic hyperplasia; antibacterial; NBD;
KM      apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;
KM      neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;
KM      keloid; gene; ss.
XX      Homo sapiens.
OS      Homo sapiens.
XX      Key      Location/Qualifiers
        CDS      1..303
        FT      /*tag= a
        FT      /product= "Human DED4 DED protein"
        FT      /note= "No start and stop codon"
        FT      /partial
XX      US2003049702-A1.
XX      13-MAR-2003.
XX      15-NOV-2001; 2001US-00001254.
XX      17-NOV-2000; 2000US-00715893.
XX      17-NOV-2000; 2000US-0367360P.
XX      29-JUN-2001; 2001US-0301889P.
XX      PA      (REED/) REED J C.
XX      PA      (GODZ/) GODZIK A.
XX      PA      (PAWL/) PAWLOWSKI K.
XX      PA      (FIOR/) FIORENTINO L.
XX      PA      (LEES/) LEE S H.
XX      PA      (ROTH/) ROTH W.
XX      PA      (STEN/) STENNER-LIEWEN F.
XX      PI      Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
XX      PI      Stenner-Liewen F;
XX      WI      MPI, 2002-500222/53.

```

```

DR      P-PSDB; AAB38898.
XX      New polypeptide comprising a death domain or death effector domain,
XX      useful for discovery of drugs that suppress infection, inflammation,
XX      allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX      Claim 18; Page 34-35; 99pp; English.
XX      The present invention provides novel death Domain (DD) and death effector
XX      domain (DED) proteins and nucleic acids encoding them. The invention also
XX      provides death domain containing protein such as Chlamydia trachomatis
XX      death domain containing protein (CTDD) DD and neural growth factor
XX      receptor-interacting death domain (NIDD) DD. The invention is useful for
XX      identifying a binding agent (e.g. protein or drug) that binds a DD, DED
XX      or NB-ARC domain from DAB3, IRAK4, CTDD, DED4 or NIDD with a candidate
XX      binding agent and identifying an effective agent (e.g. protein or drug)
XX      that modulates the association of a DD, DED or NB-ARC domain with protein
XX      that binds the DD, DED or NB-ARC domain. The invention is also useful for
XX      modulating the level of cell process such as apoptosis, cell adhesion,
XX      cell proliferation, cell stress responses, responses to microbial
XX      infection and B cell immunoglobulin class switching. DEDs and NB-ARC
XX      domains and/or anti-DD, anti-DED or anti-NB-ARC domain antibodies are
XX      useful for discovery of drugs that suppress infection, autoimmunity,
XX      inflammation, allergy, allograft rejection, sepsis and other diseases.
XX      DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
XX      autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
XX      hyperplasia, neoplasia, keloid, benign prostatic hyperplasia, fibrosis,
XX      inflammatory hyperplasia and smooth muscle cell proliferation in arteries
XX      following balloon angioplasty (restenosis). The invention is also used in
XX      antibody therapy and gene therapy. The present sequence is human DED4 DED
XX      cDNA. The DED4 gene is located on chromosome 1
XX      Sequence 303 BP; 41 A; 96 C; 115 G; 51 T; 0 U; 0 Other;
SQ      Query Match      13.4%; Score 252; DB 6; Length 303;
        Best Local Similarity 99.7%; Pred. No. 1.8e-106;
        Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      157 TGGGAGAGAGATAGTGTCTGAGCTACTACGAGATGCTGTGCTTACCGGTATGTTGAG 216
        |||
        |||
Db      1 TGGGAGAGAGATAGTGTCTGAGCTACTACGAGATGCTGTGCTTACCGGTATGTTGAG 60
QY      217 GTGGTGGGCGGCACTGACCGAGTGGAGCTGCTGCTTCTTGTCTGATGAG 276
        |||
        |||
Db      61 GTGTGGGGGGGCACTGACCGAGTGGAGCTGAGCTCTTGCTTGTCTGATGAG 120
QY      277 GCTCTGGGCGCGCCGAGAGCTTACGCCCGGCGGCTTAAGCTTCTGCTGAG 336
        |||
        |||
Db      121 GCTCTGGGCGCGCCGAGAGCTTACGCCCGGCGGCTTAAGCTTCTGCTGAG 180
QY      337 CTGAGAGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTCTGCGG 396
        |||
        |||
Db      181 CTGAGAGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTCTGCGG 240
QY      397 GTGCTGGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTCTGCT 456
        |||
        |||
Db      241 GTGCTGGCGCGCGGAGTGGCTGGCGGAGCACTGTGGGGCAACTGTCT 300
QY      457 CCA 459
        |||
        |||
Db      301 CCA 303

RESULT 26
ABL80091 standard; cDNA; 441 BP.
XX      ABL80091;
XX      17-MAY-2002 (first entry)
XX      Human ovarian cancer related cDNA clone SEQ ID NO:3069.
XX

```


KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX Homo sapiens.
XX MO200192581-A2.
XX
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001MO-US017756.
XX
XX 26-MAY-2000; 2000US-0207484P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Algarte PA, Harlocker SL, Jones R;
XX WPI; 2002-122075/16.
XX
XX Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
XX Claim 1; SEQ ID NO 3069; 489pp; English.
XX
XX The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumor
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (SI)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridizes to (SI) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridizing to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridizing to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumor protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumor polypeptides and proteins in tumor cells; and
CC to isolate a full length gene from a suitable library e.g., a tumor cDNA
CC library using well known techniques
XX
XX Sequence 441 BP; 84 A; 154 C; 113 G; 90 T; 0 U; 0 Other;
SQ
Query Match 12.6%; Score 237; DB 6; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.6e-99;
Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1209 TCTGAGGCTTAAACATTTCCAGCTGAGTTCTTCCAGACTCTTCCATCCCCAGGTTG 1268
DB 199 TCTGAGGCTTAAACATTTCCAGCTGAGTTCTTCCAGACTCTTCCATCCCCAGGTTG 258
QY 1269 TGGCCCTTATGCTCCGAGGCGGGGCTGATATTCAGAAAGGAGGCGACAGC 1328
DB 259 TGGCCCTTATGCTCCGAGGCGGGGCTGATATTCAGAAAGGAGGCGACAGC 318
QY 1329 TACACACTACCAAAAGGCGCCCTGACATTTGATCTCTGATCTTGGGCTGTGCACTG 1388
DB 319 TACACACTACCAAAAGGCGCCCTGACATTTGATCTCTGATCTTGGGCTGTGCACTG 378
QY 1389 TCACAGGTGACACACTGCTCATGTCTACACATGCCCCCTGAGATCTTCCCTGGG 1445
DB 379 TCACAGGTGACACACTGCTCATGTCTACACATGCCCCCTGAGATCTTCCCTGGG 435

AC ABL39692;
XX
XX 10-MAY-2002 (first entry)
DT
XX
XX Human NS cDNA sequence SEQ ID NO:2.
DE
XX
XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;
KW antineumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;
KW anorectic; muscular; antiferility; cardiovascular; anticoagulant;
KW antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiant;
KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
KW gastrointestinal; virucide; antilucer; cerebroprotective; nootropic;
KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;
KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;
KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KW leukaemia; asthma; immune disease; coagulation disease; hypertension;
KW diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KW gastric ulcer; Alzheimer's disease; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX MO200206315-A2.
PN
XX
XX 24-JAN-2002.
PD
XX
XX 17-JUL-2001; 2001MO-II000653.
PF
XX
XX 18-JUL-2000; 2000IL-00137345.
PR
XX
XX 15-DEC-2000; 2000IL-00140354.
XX
XX (COMP-) COMPUGEN LTD.
PA
XX
XX Mintz L, Freilich S, Bernstein J;
PI
XX
XX WPI; 2002-155037/20.
DR
XX
XX P-PADB; ABB06038.
DR
XX
XX One hundred and twenty eight novel nucleic acid sequences, useful for
PT treating and diagnosing e.g. cancer, asthma and Alzheimer's.
PT
XX
XX Claim 1; Page 76-77; 290pp; English.
PS
XX
XX ABL39691 to ABL39692 represent novel human nucleic acid sequences
CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC antineumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,
CC anorectic, muscular, anti-HIV, antiferility, cardiovascular, cardiant,
CC anticoagulant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antilucer,
CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
CC nootropic and contraceptive activities. The NS can be used in vaccines,
CC gene therapy and antisense therapy. Nucleic acids, expression vectors and
CC antibodies from the present invention can be used for treating and
CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative
CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,
CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC Alzheimer's disease and as a contraceptive
XX
XX Sequence 1084 BP; 174 A; 332 C; 378 G; 190 T; 0 U; 10 Other;
SQ
Query Match 9.9%; Score 187; DB 6; Length 1084;
Best Local Similarity 100.0%; Pred. No. 2.8e-76;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 810 GAGCGTTTGGGACGACCGCAGTGTGCTCGCTCAAGGAGACTGGGCTCTGTGTTTG 869

```

Db      764 GGAACGTTGGGCGAGCCACCGCACTGCTCCCTAAGGAGCTGGCTGTGTGGTTTG 823
Qy      870 TGAATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTCTGGGGGCACTACCTGAGTGG 929
Db      824 TGACATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTCTGGGGGCACTACCTGAGTGG 883
Qy      930 CGCCCTGCTGCAGAGCCCTTGGCGGGCGCTGTTCTGACTGAGGCGCTTCGAGAGGCTGTGGG 989
Db      884 CGCCCTGCTGCAGAGCCCTTGGCGGGCGCTGTTCTGACTGAGGCGCTTCGAGAGGCTGTGGG 943
Qy      990 CCGGGAG 996
Db      944 CCGGGAG 950

RESULT 28
ABL39693
ID      ABL39693 standard; cDNA; 1106 BP.
AC      ABL39693;
XX
XX      10-MAY-2002 (first entry)
DT
XX      Human NS cDNA sequence SEQ ID NO:3.
DE
XX
XX      Human; cytosolic; osteopathic; gynaecological; neuroprotective;
KM      antiheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;
KM      vasorectic; antiarteriosclerotic; antiinflammatory; dermatological;
KM      anorectic; muscular; antiferility; cardiovascular; anticoagulant;
KM      antifibrinolytic; hypotension; antiaesthetic; immunomodulator; cardiant;
KM      anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;
KM      gastrointestinal; virucide; antitumor; cerebroprotective; noctropic;
KM      contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;
KM      endometrios; degenerative disease; multiple sclerosis; psoriasis;
KM      rheumatoid arthritis; cataract; resection; atherosclerosis; glaucoma;
KM      inflammation; skin disorder; obesity; muscular dystrophy; AIDS;
KM      infertility; cardiovascular disease; coagulation disease; hypertension;
KM      ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;
KM      diabetes; anxiety; depression; schizophrenia; viral disease; stroke;
KM      gastric ulcer; Alzheimer's disease; gene; ss.
XX
XX      Homo sapiens.
OS
XX      WO200206315-A2.
PN
XX      24-JAN-2002.
PD
XX      17-JUL-2001; 2001WO-11000653.
PF
XX      18-JUL-2000; 2000IL-0013745.
PR      15-DEC-2000; 2000IL-00140354.
XX
XX      (COMP-) COMPUGEN LTD.
PA
XX      Mintz L, Freilich S, Bernstein J;
PI
XX      MPI; 2002-155037/20.
DR      P-PDB; ABB06039.
XX
XX      One hundred and twenty eight novel nucleic acid sequences, useful for
PT      treating and diagnosing e.g. cancer, asthma and Alzheimer's.
XX
XX      Claim 1; Page 77; 290pp; English.
PS
XX      ABL39691 to ABL39818 represent novel human nucleic acid sequences
CC      encoding the proteins given in ABB06037 to ABB06164. The novel sequences
CC      (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,
CC      antiheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,
CC      vasorectic, antiarteriosclerotic, antiinflammatory, dermatological,
CC      anorectic, muscular, anti-HIV, antiferility, cardiovascular,
CC      anticoagulant, antifibrinolytic, hypotension, antiaesthetic, cardiant,
CC      immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antitumor,

```

```

CC      antidepressant, gastrointestinal, neuroleptic, cerebroprotective,
CC      nootropic and contraceptive activities. The NS can be used in vaccines,
CC      gene therapy and antiseptic therapy. Nucleic acids, expression vectors and
CC      antibodies from the present invention can be used for treating and
CC      diagnosing e.g. cancer, osteoporosis, endometrios, degenerative
CC      diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,
CC      cataracts, resection, atherosclerosis, inflammation, skin disorder,
CC      glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular
CC      disease, coagulation disease, ischaemia, hypertension, asthma, immune
CC      disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,
CC      depression, schizophrenia, viral disease, gastric ulcers, stroke,
CC      Alzheimer's disease and as a contraceptive
XX
XX      Sequence 1106 BP; 182 A; 340 C; 385 G; 189 T; 0 U; 10 Other;
SQ
XX
XX      Query Match      9.9%; Score 187; DB 6; Length 1106;
XX      Best Local Similarity 100.0%; Pred. No. 2.8e-76;
XX      Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      810 GGAACGTTGGGCGAGCCACCGCACTGCTCCCTAAGGAGCTGGCTGTGTGGTTTG 869
Db      786 GGAACGTTGGGCGAGCCACCGCACTGCTCCCTAAGGAGCTGGCTGTGTGGTTTG 845
Qy      870 TGAATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTCTGGGGGCACTACCTGAGTGG 929
Db      846 TGACATCAAGTTCTCAGAGCTCTCTATCTGACGCGCTTCTGGGGGCACTACCTGAGTGG 905
Qy      930 CGCCCTGCTGCAGAGCCCTTGGCGGGCGCTGTTCTGACTGAGGCGCTTCGAGAGGCTGTGGG 989
Db      906 CGCCCTGCTGCAGAGCCCTTGGCGGGCGCTGTTCTGACTGAGGCGCTTCGAGAGGCTGTGGG 965
Qy      990 CCGGGAG 996
Db      966 CCGGGAG 972

RESULT 29
AAH12893/C
ID      AAH12893 standard; cDNA; 381 BP.
AC      AAH12893;
XX
XX      26-JUN-2001 (first entry)
DT
XX
XX      Human cDNA clone (3'-primer) SEQ ID NO:9728.
DE
XX
XX      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX      Homo sapiens.
OS
XX      EP1074617-A2.
PN
XX      EP1074617-A2.
PD
XX      07-FEB-2001.
PF
XX      28-JUL-2000; 2000EP-00116126.
PR      29-JUL-1999; 99JP-00248036.
PR      27-AUG-1999; 99JP-00300253.
PR      11-JAN-2000; 2000JP-00118776.
PR      02-MAY-2000; 2000JP-001183767.
PR      09-JUN-2000; 2000JP-00241899.
XX
XX      (HELI-) HELIX RES INST.
PA
XX      Oca T, Isogai T, Nishikawa T, Hayaehi K, Saito K, Yamamoto J;
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX      MPI; 2001-318749/34.
DR
XX
XX      Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT      length cDNAs defined in the specification, and for the detection and/or
PT      diagnosis of the abnormality of the proteins encoded by the full-length
PT      cDNAs.

```

XX	PS	Claim 3; SEQ ID NO 9728; 2537bp + Sequence Listing; English.
CC	CC	The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
CC	CC	(a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602
CC	CC	nucleotide sequences defined in the specification, where the
CC	CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC	CC	of an oligonucleotide comprising a sequence complementary to the
CC	CC	complementary strand of a polynucleotide which comprises a 5'-end
CC	CC	sequence and an oligonucleotide comprising a sequence complementary to a
CC	CC	polynucleotide which comprises a 3'-end sequence, where the
CC	CC	oligonucleotide comprises at least 15 nucleotides and the combination of
CC	CC	the 5'-end sequence/3'-end sequence is selected from those defined in the
CC	CC	specification. The primer sets can be used in antisense therapy and in
CC	CC	gene therapy. The primers are useful for synthesising polynucleotides,
CC	CC	particularly full-length cDNAs. The primers are also useful for the
CC	CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	CC	oligonucleotides, all of which are used in the exemplification of the
CC	CC	present invention
XX	XX	
SQ	SQ	Sequence 381 BP; 84 A; 95 C; 88 G; 100 T; 0 U; 14 Other;
QY	Query Match	8.8%; Score 165; DB 4; Length 381;
	Best Local Similarity	99.5%; Pred. No. 4.8e-66;
	Matches	215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1668	TGTCAGAGCTGAACATAGACTTGACCTTGATGTCACTGAGTGCATGTGGAGTGT 17277
DB	216	TGTCAGAGCTGAACATAGACTTGATGTCACTGAGTGCATGTGGAGTGT 157
QY	1728	TATGGCAGCATCATACCAAGGCTTACTTTTGACATGTGGGGCCAAAACCATGTAACAGCCA 1787
DB	156	TATGGCAGCATCATACCAAGGCTTACTTTTGACATGTGGGGCCAAAACCATGTAACAGCCA 97
QY	1788	CCTTCTTGGAAGGGAATGCAAGGCTTTGGGGGATGTAAGAAAGCCTTTTACAATGCA 1847
DB	96	CCTTCTTGGAAGGGAATGCAAGGCTTTGGGGGATGTAAGAAAGCCTTTTACAATGCA 37
QY	1848	TACCAATTAACTGCCCTGTGAAGGGCATAGTGTGGG 1883
DB	36	TACCAATTAACTGCCCTGTGAAGGGCATAGTGTGGG 1
	RESULT 30	
	ACH25056	
	ID	ACH25056 standard; cDNA; 484 BP.
	XX	
	AC	ACH25056;
	XX	
	DT	13-OCT-2003 (first entry)
	XX	
	DE	Human adult ovary cDNA #3436.
	XX	
	OS	Homo sapiens.
	XX	
	FN	US2003073623-A1.
	XX	
	PD	17-APR-2003.
	XX	
	PF	30-JUL-2001; 2001US-00918995.
	XX	
	PR	30-JUL-2001; 2001US-00918995.
	XX	
	PA	(DRMA/) DRMANAC R T.

PA	(LBA// LABAT I.
PA	(STAC// STACHE-CRAIN B.
PA	(DICK// DICKSON M C.
PA	(JONE// JONES L W.
PI	Drmanc RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
DR	WPI; 2003-615964/58.
XX	
XX	New polynucleotide sequences obtained from various cDNA libraries, useful
PT	as hybridization probes, as oligomers for PCR, for chromosome and gene
PT	mapping, in the recombinant production of protein, or in generating
PT	antisense DNA or RNA.
XX	
PS	Claim 1; SEQ ID NO 12268; 44pp; English.
CC	The invention relates to an isolated polynucleotide comprising any one of
CC	38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC	determined by the technique of SSH (sequencing by hybridisation). Also
CC	included is a purified polypeptide comprising a sequence corresponding to
CC	a reading frame of the novel polynucleotide. The nucleic acid sequences
CC	are useful in diagnostics as expressed sequence tags (EST) for
CC	identifying expressed genes or for physical mapping of the human genome,
CC	in forensics, in assessing biodiversity, or in identifying mutations
CC	responsible for genetic disorders and other traits. The nucleotide
CC	sequences are also useful as hybridisation probes, as oligomers for PCR,
CC	for chromosome and gene mapping, in the recombinant production of
CC	protein, or in generating antisense DNA or RNA. The purified polypeptide
CC	is useful for generating antibodies specific for it. The present sequence
CC	is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC	for this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from USPTO at
CC	seqdata.uspto.gov/sequence.html?docID=20030073623
XX	
SQ	Sequence 484 BP; 83 A; 148 C; 159 G; 93 T; 0 U; 1 Other;
XX	
Query Match	5.1%; Score 96; DB 9; Length 484;
Best Local Similarity	100.0%; Pred. No. 5.6e-34;
Matches	96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	873 CATCAAGTTCGAAGCTCCTCATCTCGAAGCGCTTGCGGGGCGACTAACGTAGTGGCGC 932
Db	36 CATCAAGTTCGAAGCTCCTCATCTCGAAGCGCTTGCGGGGCGACTAACGTAGTGGCGC 95
OY	933 CCTGCTGAGGCGCTTGCGGGGCGTGTCTCGACTGA 968
Db	96 CCTGCTGAGGCGCTTGCGGGGCGTGTCTCGACTGA 131
RESULT 31	
ADCT9268	ID ADCT9268 standard; DNA; 45 BP.
XX	
AC	ADCT9268;
XX	
DT	01-JAN-2004 (first entry)
XX	
DE	Human DEDD2 PCR primer SEQ ID NO:10.
XX	
KM	human; death effector domains containing DNA-binding protein;
KM	DEO-containing DNA-binding protein; DEDD2; cell death; gene therapy;
KM	cyclostatic; cancer; chronic myeloid leukaemia; PCR primer; ss.
XX	
OS	Synthetic.
XX	
OS	Homo sapiens.
XX	
PN	WO2003054195-A1.
XX	
PD	03-JUL-2003.
XX	
PF	20-DEC-2002; 2002WO-JP013371.
XX	
PR	20-DEC-2001; 2001JP-00387854.

```

PR 18-JUL-2002; 2002JP-00209458.
XX (MORG ) MORINAGA MILK IND CO LTD.
XX
XX
PI Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX WPI, 2003-569246/53.
XX
XX DNA encoding cell death proteins for treatment of kidney, large intestine
PT and prostate cancers and leukemia.
XX
XX Example; Page 23; 26pp; Japanese.
XX
XX The present invention describes a human death effector domains (DED)
CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that
CC causes cell death. Also described: (1) primer and probe for investigation
CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
CC cytostatic activity. DEDD2 can be used in the diagnosis and treatment of
CC cancers of the kidney, large intestine and prostate, and acute and
CC chronic myeloid leukaemia. The present sequence represents a PCR primer
CC which is used in an example from the present invention.
XX
XX Sequence 45 BP; 9 A; 17 C; 11 G; 8 T; 0 U; 0 Other;
SQ
Query Match 2.4%; Score 45; DB 10; Length 45;
Best Local Similarity 100.0%; Pred. No. 3.1e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1105 TCCAGAGCTGGCAGATTGATCCACCTCCAGCTCCGGGCGAC 1149
DB 1 TCCAGAGCTGGCAGATTGATCCACCTCCAGCTCCGGGCGAC 45
RESULT 32
ID ABZ70271 standard; DNA; 41 BP.
XX
XX ABZ70271;
AC
XX
XX 25-APR-2003 (first entry)
DT
XX
XX ATP synthetase 11.55 probe #1.
DE
XX
XX ATP synthetase 11.55; enzyme; cancer; cytostatic; HIV infection;
KW anti-HIV; probe; ss.
XX
XX Unidentified.
OS
XX
XX CN1363656-A.
PN
XX
XX 14-AUG-2002.
PD
XX
XX 05-JAN-2001; 2001CN-00105031.
PF
XX
XX 05-JAN-2001; 2001CN-00105031.
PR
XX
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.
PA
XX
XX Mao Y, Xie Y;
PI
XX
XX WPI, 2002-751770/82.
DR
XX
XX Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.
PT
XX
XX Example 6; Page 19 (Disclosure); 32pp; Chinese.
PS
XX
XX The present invention relates to ATP synthetase 11.55 (see ABP59171). The
CC protein can be used for treating diseases such as cancer and HIV
CC infection. The present sequence is a probe, which was used in an example
CC from the invention
XX
XX Sequence 41 BP; 7 A; 11 C; 10 G; 13 T; 0 U; 0 Other;
SQ

```

```

Query Match 2.2%; Score 41; DB 6; Length 41;
Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1202 TGTGGGCTCTGAGGCGCTTAACCATTTCCAGCTGAGTTTCT 1242
DB 1 TGTGGGCTCTGAGGCGCTTAACCATTTCCAGCTGAGTTTCT 41
RESULT 33
ID AA158961/C
XX AA158961 standard; cDNA; 1067 BP.
XX
XX AA158961;
AC
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 1164.
DE
XX
XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US034263.
PF
XX
XX 23-DEC-1999; 99US-00471275.
PR
XX
XX 25-APR-2000; 2000US-00498725.
PR
XX
XX 20-JUN-2000; 2000US-00582317.
PR
XX
XX 19-JUL-2000; 2000US-00598042.
PR
XX
XX 03-AUG-2000; 2000US-00620312.
PR
XX
XX 14-SEP-2000; 2000US-00653450.
PR
XX
XX 19-OCT-2000; 2000US-00662191.
PR
XX
XX 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR
XX
XX P-PADB; AAM39805.
DR
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 1164; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AA42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

```

XX SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;
Query Match 2.0%; Score 37; DB 4; Length 1067;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1760 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 1796
DB 979 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 943
RESULT 34
ADQ9183/c
ID ADQ9183 standard; cDNA; 1067 BP.
XX
XX ADQ9183;
AC
XX 23-SEP-2004 (first entry)
DT
XX
XX DNA encoding human GPCR-like protein segid 853.
DE
XX
XX ophthalmological; immunomodulatory; cytostatic; antiatherosclerotic;
KM antidiabetic; GPCR-like protein; ophthalmic disorder;
KM neurological disorder; immunological disorder; nephritic disorder;
KM hormonal dysfunction; cancer; atherosclerosis; diabetes;
KM molecular weight marker; food supplement; human; ss.
XX
XX Homo sapiens.
XX
XX US6569662-B1.
PN
XX
XX 27-MAY-2003.
PD
XX
XX 19-JUL-2000; 2000US-00620312.
PF
XX
XX 21-JAN-2000; 2000US-00488725.
PR
XX 25-APR-2000; 2000US-00552317.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Zhou P, Drmanac RT;
PI
XX
XX WPI; 2001-442255/47.
DR
XX
XX New G-protein-coupled receptor-like polypeptides and polynucleotides,
PT useful for treating diseases of ophthalmic, neurological, immunological
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis
PT and diabetes.
XX
XX Example 2; SEQ ID NO 853; 92pp; English.
PS
XX
XX The invention describes an isolated polynucleotide (I) comprising a fully
CC defined (S1) of 745, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041,
CC 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as
CC given in the specification, its translated or protein coding portion, its
CC extracellular portion or its active domain. The GPCR-like polypeptides
CC and polynucleotides are useful for the treatment of diseases of
CC ophthalmic, neurological, immunological and nephritic systems. They may
CC also be used to treat hormonal dysfunction, cancer, atherosclerosis and
CC diabetes. The antibodies are useful for detecting or quantitating the
CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.
XX
XX
SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;
Query Match 2.0%; Score 37; DB 5; Length 1067;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1760 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 1796
DB 979 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 943

DB 979 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 943
RESULT 35
ADB48943/c
ID ADB48943 standard; cDNA; 1067 BP.
XX
XX ADB48943;
AC
XX
XX 04-DEC-2003 (first entry)
DT
XX
XX Novel human cDNA SEQ ID NO 853.
DE
XX
XX ss; cancer; neurodegenerative disease; human.
KM
XX
XX Homo sapiens.
OS
XX
XX US2003104529-A1.
PN
XX
XX 05-JUN-2003.
PD
XX
XX 04-JAN-2002; 2002US-00037270.
PF
XX
XX 21-JAN-2000; 2000US-00488725.
PR
XX 25-APR-2000; 2000US-00552317.
XX
XX 19-JUL-2000; 2000US-00620312.
XX
XX (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIUC/) LIU C.
PA (ASUNDI/) ASUNDI V.
PA (DRMA/) DRMANAC R T.
XX
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
XX
XX WPI; 2003-678194/64.
DR
XX
XX New polynucleotide, useful for treating diseases e.g., cancer or
PT neurodegenerative diseases.
PT
XX
XX Claim 1; SEQ ID NO 853; 99pp; English.
PS
XX
XX The invention relates to a polynucleotide comprising a sequence given in
CC the specification, or its mature protein-coding portion, or its
CC complement. The polynucleotide is useful for treating diseases e.g.,
CC cancer or neurodegenerative diseases and many others listed in the
CC specification. The present sequence represents a novel human cDNA. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docid=20030104529.
XX
XX
SQ Sequence 1067 BP; 191 A; 349 C; 328 G; 199 T; 0 U; 0 Other;
Query Match 2.0%; Score 37; DB 9; Length 1067;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1760 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 1796
DB 979 ACATGGGGCCAAAACAGTAAACAGCCACCTTCTTG 943
RESULT 36
ADQ55421
ID ADQ55421 standard; DNA; 594 BP.
XX
XX ADQ55421;
AC
XX
XX 21-OCT-2004 (first entry)
DT
XX
XX Novel canine microarray-related DNA sequence SegID6723.
DE
XX
XX canine microarray; drug screening; toxicity assay;
KM

KW environmental pollutant; cellular response; gene expression profile;
KM toxic response; liver necrosis; fatty liver disease;
XX protein adduct formation; hepatitis; dog; da.
OS
XX
XX Canis familiaris.
PN WO2004063324-A2.
PP
PD 29-JUL-2004.
XX
PF 05-MAY-2003; 2003WO-US013853.
XX
PR 03-MAY-2002; 2002US-0377240P.
XX
PA (GENE-) GENE LOGIC INC.
XX (Pfiz) PFIZER PROD INC.
PI Digians JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
DR
XX
XX New isolated nucleic acid molecule, useful for drug screening and
PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX
XX Claim 1; SEQ ID NO 6723; 41pp; English.
PS
XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northernblots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific
CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least
CC one of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.
XX
SQ Sequence 594 BP; 124 A; 193 C; 141 G; 120 T; 0 U; 16 Other;

Query Match 1.5%; Score 29; DB 13; Length 594;
Best Local Similarity 100.0%; Pred.No. 0.0078;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1575 GGGGCGTGTCTACCTTGCCACACATTC 1603
DB 312 GGGGCGTGTCTACCTTGCCACACATTC 340

RESULT 37
ADC79267
ID ADC79267 standard; DNA; 26 BP.
XX
AC ADC79267;

XX	
DT	01-JAN-2004 (first entry)
XX	
DE	Human DEDD2 PCR primer SEQ ID NO:9.
XX	
KW	human; death effector domains containing DNA-binding protein;
KM	DED-containing DNA-binding protein; DEDD2; cell death; gene therapy;
XX	cystostatic; cancer; chronic myeloid leukaemia; PCR primer; ss.
XX	
OS	Synthetic.
XX	Homo sapiens.
XX	
PV	WO2003054195-A1.
XX	
PD	03-JUL-2003.
XX	
PF	20-DEC-2002; 2002WO-JP013371.
XX	
PR	20-DEC-2001; 2001JP-00387854.
XX	18-JUL-2002; 2002JP-00209458.
PA	(MORG) MORINAGA MILK IND CO LTD.
XX	
PI	Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX	
DR	WPI; 2003-569246/53.
XX	
PT	DNA encoding cell death proteins for treatment of kidney, large intestine
XX	and prostate cancers and leukemia.
XX	
PS	Example; Page 23; 26pp; Japanese.
XX	
CC	The present invention describes a human death effector domains (DED)
CC	containing DNA-binding protein (DEDD) protein, designated DEDD2, that
CC	causes cell death. Also described: (1) primer and probe for investigation
CC	of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
CC	cyclostetic activity. DEDD2 can be used in the diagnosis and treatment of
CC	cancers of the kidney, large intestine and prostate, and acute and
CC	chronic myeloid leukaemia. The present sequence represents a PCR primer
CC	which is used in an example from the present invention.
XX	
XX	
SO	Sequence 26 BP; 1 A; 13 C; 6 G; 6 T; 0 U; 0 Other;
	Query Match 1.4%; Score 26; DB 10; Length 26;
	Best Local Similarity 100.0%; Pred. No. 0.21;
	Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	98 GCTGTTCGGCCTCCCTCCCGGGA 123
DB	1 GCTGTTCGGCCTCCCTCCCGGGA 26
	RESULT 38
ID	ADC79264/C
XX	ADC79264 standard; DNA; 25 BP.
XX	
AC	ADC79264;
XX	
DT	01-JAN-2004 (first entry)
XX	
DE	Human DEDD2 related PCR primer SEQ ID NO:6.
XX	
KW	human; death effector domains containing DNA-binding protein;
KM	DED-containing DNA-binding protein; DEDD2; cell death; gene therapy;
XX	cystostatic; cancer; chronic myeloid leukaemia; PCR primer; ss.
XX	
OS	Synthetic.
XX	Homo sapiens.
XX	
PN	WO2003054195-A1.
XX	
PD	03-JUL-2003.
XX	

```

PF 20-DEC-2002; 2002MO-JP013371.
XX
XX 20-DEC-2001; 2001JP-00387854.
PR 18-JUL-2002; 2002JP-00209458.
XX
XX (MORG ) MORINAGA MILK IND CO LTD.
PI Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX WPI; 2003-569246/53.
XX
XX DNA encoding cell death proteins for treatment of kidney, large intestine
PT and prostate cancers and leukemia.
XX
XX Example; Page 22; 26pp; Japanese.
XX
XX The present invention describes a human death effector domains (DED)
CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that
CC causes cell death. Also described: (1) primer and probe for investigation
CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
CC cytosstatic activity. DEDD2 can be used in the diagnosis and treatment of
CC cancers of the kidney, large intestine and prostate, and acute and
CC chronic myeloid leukemia. The present sequence represents a PCR primer
CC which is used in an example from the present invention.
XX
XX Sequence 25 BP; 5 A; 3 C; 11 G; 6 T; 0 U; 0 Other:
SQ
Query Match 1.3%; Score 25; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1398 CACACACTCGCTCATGCTCACACTG 1422
Db 25 CACACACTCGCTCATGCTCACACTG 1
XX
XX RESULT 39
XX ADCT9263
XX ID ADC79263 standard; DNA; 25 BP.
XX
XX ADC79263;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human DEDD2 related PCR primer SEQ ID NO:5.
XX
XX human; death effector domains containing DNA-binding protein;
XX DED-containing DNA-binding protein; DEDD2; cell death; gene therapy;
XX cytosstatic; cancer; chronic myeloid leukemia; PCR primer; ss.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO2003054195-A1.
XX
XX 03-JUL-2003.
XX
XX 20-DEC-2002; 2002MO-JP013371.
XX
XX 20-DEC-2001; 2001JP-00387854.
PR 18-JUL-2002; 2002JP-00209458.
XX
XX (MORG ) MORINAGA MILK IND CO LTD.
PA
XX Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX WPI; 2003-569246/53.
XX
XX DNA encoding cell death proteins for treatment of kidney, large intestine
PT and prostate cancers and leukemia.
XX
XX Example; Page 22; 26pp; Japanese.
XX

```

```

CC The present invention describes a human death effector domains (DED)
CC containing DNA-binding protein (DEDD) protein, designated DEDD2, that
CC causes cell death. Also described: (1) primer and probe for investigation
CC of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
CC cytosstatic activity. DEDD2 can be used in the diagnosis and treatment of
CC cancers of the kidney, large intestine and prostate, and acute and
CC chronic myeloid leukemia. The present sequence represents a PCR primer
CC which is used in an example from the present invention.
XX
XX Sequence 25 BP; 5 A; 9 C; 6 G; 5 T; 0 U; 0 Other:
SQ
Query Match 1.3%; Score 25; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 444 CCGGCCAGTGTCTCCAGAACGCTAT 468
Db 1 CCGGCCAGTGTCTCCAGAACGCTAT 25
XX
XX RESULT 40
XX ABZ70269
XX ID ABZ70269 standard; DNA; 33 BP.
XX
XX ABZ70269;
XX
XX 25-APR-2003 (first entry)
XX
XX ATP synthetase 11.55 PCR primer #3.
DE
XX ATP synthetase 11.55; enzyme; cancer; cytosstatic; HIV infection;
XX anti-HIV; PCR; primer; ss.
XX
XX Unidentified.
XX
XX CN1363656-A.
XX
XX 14-AUG-2002.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX 05-JAN-2001; 2001CN-00105031.
PR
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.
PA
XX Mao Y, Xie Y;
PI
XX WPI; 2002-751770/82.
XX
XX Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.
PT
XX Example 4; Page 17 (Disclosure); 32pp; Chinese.
XX
XX The present invention relates to ATP synthetase 11.55 (see ABP59171). The
CC protein can be used for treating diseases such as cancer and HIV
CC infection. The present sequence is a PCR primer, which was used in an
CC example from the invention
XX
XX Sequence 33 BP; 8 A; 10 C; 8 G; 7 T; 0 U; 0 Other:
SQ
Query Match 1.3%; Score 25; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1200 GATGTGGCTCTGAGGCTTAACCA 1224
Db 9 GATGTGGCTCTGAGGCTTAACCA 33
XX
XX RESULT 41
XX ABZ70267
XX ID ABZ70267 standard; DNA; 24 BP.
XX

```

```

AC ABZ70267;
XX
XX 25-APR-2003 (first entry)
XX
XX ATP synthetase 11.55 PCR primer #1.
DE
XX ATP synthetase 11.55; enzyme; cancer; cytostatic; HIV infection;
XX anti-HIV; PCR; primer; ss.
XX
XX Unidentified.
XX
XX CN1363656-A.
XX
XX 14-AUG-2002.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2002-751770/82.
XX
XX Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.
XX
XX Example 2; Page 17 (Disclosure); 32pp; Chinese.
XX
XX The present invention relates to ATP synthetase 11.55 (see ABP59171). The
XX protein can be used for treating diseases such as cancer and HIV
XX infection. The present sequence is a PCR primer, which was used in an
XX example from the invention
XX
XX Sequence 24 BP; 4 A; 11 C; 5 G; 4 T; 0 U; 0 Other;
SQ
Query Match 1.3%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1123 GATCCACCTCCAGTCTCGGGC 1146
DB 1 GATCCACCTCCAGTCTCGGGC 24
RESULT 42
ABZ70270/C
ID ABZ70270 standard; DNA; 33 BP.
XX
XX ABZ70270;
XX
XX 25-APR-2003 (first entry)
XX
XX ATP synthetase 11.55 PCR primer #4.
XX
XX ATP synthetase 11.55; enzyme; cancer; cytostatic; HIV infection;
XX anti-HIV; PCR; primer; ss.
XX
XX Unidentified.
XX
XX CN1363656-A.
XX
XX 14-AUG-2002.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX 05-JAN-2001; 2001CN-00105031.
XX
XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2002-751770/82.

```

```

XX Polypeptide-ATP synthetase 11.55 and polynucleotide for coding it.
XX
XX Example 4; Page 17 (Disclosure); 32pp; Chinese.
XX
XX The present invention relates to ATP synthetase 11.55 (see ABP59171). The
XX protein can be used for treating diseases such as cancer and HIV
XX infection. The present sequence is a PCR primer, which was used in an
XX example from the invention
XX
XX Sequence 33 BP; 11 A; 6 C; 9 G; 7 T; 0 U; 0 Other;
SQ
Query Match 1.3%; Score 24; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1495 TCTCTCTCAGACCTCTAATTGA 1518
DB 33 TCTCTCTCAGACCTCTAATTGA 10
RESULT 43
ADC79265
ID ADC79265 standard; DNA; 23 BP.
XX
XX ADC79265;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human DEDD2 PCR primer SEQ ID NO:7.
XX
XX human; death effector domains containing DNA-binding protein;
XX DED-containing DNA-binding protein; DEDD2; cell death; gene therapy;
XX cytostatic; cancer; chronic myeloid leukemia; PCR primer; ss.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX WO2003054195-A1.
XX
XX 03-JUL-2003.
XX
XX 20-DEC-2002; 2002WO-JP013371.
XX
XX 20-DEC-2001; 2001JP-00387854.
XX
XX 18-JUL-2002; 2002JP-00209458.
XX
XX (MORG) MORINAGA MILK IND CO LTD.
XX
XX Suzu S, Hageshita H, Nomaguchi K, Yamada M, Hayasawa H;
XX
XX WPI; 2003-569246/53.
XX
XX DNA encoding cell death proteins for treatment of kidney, large intestine
XX and prostate cancers and leukemia.
XX
XX Claim 6; Page 22; 26pp; Japanese.
XX
XX The present invention describes a human death effector domain (DED)
XX containing DNA-binding protein (DEDD) protein, designated DEDD2, that
XX causes cell death. Also described: (1) primer and probe for investigation
XX of the DEDD2 gene; and (2) reagents for gene therapy. DEDD2 has
XX cytostatic activity. DEDD2 can be used in the diagnosis and treatment of
XX cancers of the kidney, large intestine and prostate, and acute and
XX chronic myeloid leukemia. The present sequence represents a PCR primer
XX which is used in an example from the present invention.
XX
XX Sequence 23 BP; 0 A; 13 C; 4 G; 6 T; 0 U; 0 Other;
SQ
Query Match 1.2%; Score 23; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

ID ABL88384 standard; cDNA; 1339 BP.
XX
AC ABL88384;
XX
XX
DT 16-MAY-2002 (first entry)
XX
DE Pain regulated cDNA sequence 27.
XX
XX Pain; analgesic; gene therapy; neurological disorder;
KW neurodegenerative disease; gene; ss.
XX
OS Rattus sp.
XX
PN WO200212338-A2.
XX
PD 14-FEB-2002.
XX
PF 03-AUG-2001; 2001WO-EP009011.
XX
PR 03-AUG-2000; 2000DE-01037759.
XX
PA (CHEF ) GRUENENTHAL GMBH.
XX
PI Gillen C, Wetzel S, Wnendt S, Weihe E, Schaefer MK;
XX
DR WPI; 2002-257469/30.
XX
XX Identifying pain-regulating compounds, useful for treating chronic pain
PT and for diagnosis, by measuring binding of compounds to specific peptides
PT and proteins.
XX
PS Claim 14; Fig 14; 213pp; German.
XX
XX The invention relates to identifying pain-regulating substances (A)
CC comprises (i) incubating a test substance with a cell (or preparation
CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
CC either binding of the test substance to (B) or some functional parameter
CC that is altered by this binding. The method is useful for identifying
CC pain-regulating substances (A) with analgesic activity. (A) along with
CC nucleic acid (ABL88411-ABL88441) that encode proteins (B, ABB85006-
CC ABB85037) that interact with (A); (B); vectors containing the nucleic
CC acid; antibodies against (B); cells that express (B) and agents that bind
CC to (B); are all useful for treating pain, particularly chronic pain,
CC including use in gene therapy. The same materials can also be used for
CC diagnosis, e.g. of neurological and neurodegenerative diseases. The
CC present sequence is that of a polynucleotide of the invention
XX
SQ Sequence 1339 BP; 385 A; 299 C; 274 G; 377 T; 0 U; 4 Other;
XX
Query Match 1.2%; Score 22; DB 6; Length 1339;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1449 CTGCCCTGGCGCTGCTCCGAGC 1470
DB 1018 CTGCCCTGGCGCTGCTCCGAGC 1039
XX
RESULT 47
AD063628
ID AD063628 standard; cDNA; 2718 BP.
XX
AC AD063628;
XX
XX
DT 07-OCT-2004 (first entry)
XX
XX Novel human cDNA sequence #789.
XX
XX
XX sb; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX

```

```

OS Homo sapiens.
XX
XX EP1440981-A2.
XX
XX
XX
PD 28-JUL-2004.
XX
XX
PF 21-JAN-2004; 2004EP-00001196.
XX
XX
XX 21-JAN-2003; 2003JP-00102206.
PR 09-MAY-2003; 2003JP-00131392.
XX
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX
XX WPI; 2004-535376/52.
DR P-FSDB; AD065816.
XX
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
XX Claim 1; SEQ ID NO 789; 2449pp; English.
XX
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 2718 BP; 575 A; 864 C; 651 G; 628 T; 0 U; 0 Other;
XX
Query Match 1.2%; Score 22; DB 12; Length 2718;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1262 CCAGGTGTCCTTACGCTC 1283
DB 870 CCAGGTGTCCTTACGCTC 891
XX
RESULT 48
AAF27413
ID AAF27413 standard; DNA; 21 BP.
XX
AC AAF27413;
XX
XX
DT 24-APR-2001 (first entry)
XX
XX Human apoptosis-associated factor NT2RM1000558 PCR primer. SEQ ID NO:10.
XX
XX Human; apoptosis-associated factor; NT2RM1000558; death effector domain;
KW DED; caspase family cleavage domain; pro-apoptotic; drug screening;
KW cell proliferation; ischaemic disease; chronic viral disease; PCR primer;
KW ss.
XX
XX Homo sapiens.
OS
XX
XX WO200104300-A1.
PN
XX
XX 18-JAN-2001.
PD
XX
XX 06-JUL-2000; 2000WO-JP004516.
PF
XX
XX 08-JUL-1999; 99JP-00194179.
PR 18-OCT-1999; 99US-0159586P.
XX
XX (HELI-) HELIX RES INST.
PA
XX

```

PI Oca T, Isogai T, Nishikawa T, Kawai Y, Miyoshi S, Satoh S;
XX WPI; 2001-138348/14.
XX
PT Polynucleotide encoding an apoptosis-associated factor protein with death
XX effector domain and caspase family-cleavage domain, useful in regulating
PT diseases with cell proliferation.
XX
PS Example 4; Page 32; 53pp; Japanese.
XX
CC The invention relates to a novel human apoptosis-associated factor
CC (AAB60386, AAB60387), designated NTZRM1000558, which contains a death
CC effector domain (DED) and a caspase family cleavage domain and is capable
CC of inducing apoptosis in cells. The invention also relates to nucleic
CC acids encoding the protein (AAF27407, AAF27408); variants of the protein
CC (particularly dominant negative variants); vectors and host cells
CC comprising a nucleic acid which encodes an apoptosis-associated factor
CC of the invention; the recombinant production of the protein; an antibody
CC against the protein; and methods of screening for compounds which can
CC regulate apoptosis. The apoptosis-related factor is useful in regulating
CC diseases associated with cell proliferation and in screening drug
CC candidates e.g., for regulating cell proliferation or cell death in
CC ischemic diseases and chronic viral diseases. The present sequence
CC represents a PCR primer used to amplify cDNA encoding the human apoptosis
CC-associated factor NTZRM1000558
XX
SQ Sequence 21 BP; 4 A; 3 C; 8 G; 6 T; 0 U; 0 Other;
Query Match 1.1%; Score 21; DB 5; Length 21;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1013 TCAGTGTGATGAGGTGACT 1033
DB 1 TCAGTGTGATGAGGTGACT 21
RESULT 49
AAD40093
ID AAD40093 standard; DNA; 21 BP.
XX
AC AAD40093;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human DED4 (death effector domain) cDNA amplifying primer #2.
XX
KW Human; death domain; DD; death effector domain; DED; Chlamydia infection;
KW NB-ARC domain; apoptosis; oncogenic protein; bacterial infection; sepsis;
KW inflammation; allergy; autoimmunity; allograft rejection; cell division;
KW immune-based pathology; fibrosis; arthritis; graft versus host disease;
KW immunosuppressive; gene therapy; antisense therapy; primer; ss.
XX
OS Homo sapiens.
XX
PN WO200240680-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001MO-US044844.
XX
PR 17-NOV-2000; 2000US-00715893.
XX
PR 29-JUN-2001; 2001US-0301889P.
XX
PA (BURN-) BURHAM INST.
XX
PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
XX Stenner-Liewen F;
XX WPI; 2002-500222/53.
XX
PT New polypeptide comprising a death domain or death effector domain,
PT useful for discovery of drugs that suppress infection, inflammation,

PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.
XX
PS Example 7; Page 118; 209pp; English.
XX
CC The invention relates to an isolated polypeptide comprising a death
CC domain (DD), death effector domain (DED) or NB-ARC domain. The invention
CC is useful for identifying a binding agent, preferably a protein or a drug
CC that binds a DD, DED or NB-ARC domain, by contacting a DD, DED or NB-ARC
CC domain from DAP3, IRAK4, CTDP (Chlamydia trachomatis DD protein), DED4 or
CC NIDD (NGFR-interacting Death Domain), with a candidate binding agent and
CC detecting the association of the domain and the candidate binding agent,
CC by yeast two hybrid assay, immunoprecipitation, SPA, ultraviolet (UV) or
CC chemical crosslinking, nuclear magnetic resonance (NMR), mass
CC spectroscopy (MS) and FPA. The invention is useful for modulating the
CC level of a cell process such as cell proliferation, cell adhesion, cell
CC stress responses, responses to microbial infection and B cell
CC immunoglobulin class switching, in particular apoptosis within a cell.
CC Antibody specifically reactive with CTDP DD of C. trachomatis, C.
CC muridarum, C. pneumoniae, and C. psittaci or a nucleic acid encoding the
CC CTDP DD protein is useful for detecting a Chlamydia infection. The
CC invention is useful for modulating the activity of oncogenic proteins,
CC for treating a pathology caused by the oncogenic proteins and for
CC treating bacterial infections by modulating the activity of bacterial
CC proteins. The protein and antibody specific for it are useful for
CC discovery of drugs that suppress infection, inflammation, allergy,
CC sepsis, autoimmunity, allograft rejection and other diseases. The protein
CC is useful for treating immune-based pathologies, pathologies associated
CC with cell division, inflammatory diseases such as sepsis, fibrosis,
CC arthritis, graft versus host disease. The invention is used in antisense
CC therapy and gene therapy. The present sequence is human DED4 cDNA
XX
SQ Sequence 21 BP; 4 A; 8 C; 7 G; 2 T; 0 U; 0 Other;
Query Match 1.1%; Score 21; DB 6; Length 21;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 56 CGCCGAGAGTCACTCGAGCC 76
DB 1 CGCCGAGAGTCACTCGAGCC 21
RESULT 50
AAD59075
ID AAD59075 standard; DNA; 21 BP.
XX
AC AAD59075;
XX
DT 18-DEC-2003 (first entry)
XX
DE Primer #2 used to amplify human DED4 cDNA.
XX
KW Human; death domain; DD; death effector domain; DED; cell proliferation;
KW Chlamydia trachomatis death domain containing protein; fibrosis; sepsis;
KW neural growth factor receptor-interacting death domain; cell adhesion;
KW vasotropic; microbial infection; inflammation; allograft rejection; CTDP;
KW cell stress response; benign prostatic hyperplasia; antibacterial; NIDD;
KW apoptosis; infection; autoimmunity; allergy; hyperplasia; gene therapy;
KW neoplasia; restenosis; immunosuppressive; antibody therapy; cytostatic;
KW keloid; primer; ss.
XX
OS Homo sapiens.
XX
PN US2003049702-A1.
XX
PD 13-MAR-2003.
XX
PF 15-NOV-2001; 2001US-00001254.
XX
PR 17-NOV-2000; 2000US-00715893.
XX
PR 17-NOV-2000; 2000US-0367360P.
XX
PR 29-JUN-2001; 2001US-0301889P.

XX (REED/) REED J C.
 PA (GODZ/) GODZIK A.
 PA (PAWL/) PAWLOWSKI K.
 PA (FIOR/) FIORENTINO L.
 PA (LEES/) LEE S H.
 PA (ROTH/) ROTH W.
 PA (STEN/) STENNER-LIEWEN F.

XX
 PI Reed JC, Godzik A, Pawlowski K, Fiorentino L, Lee SH, Roth W;
 PI Stenner-Liewen F;

XX
 DR WPI; 2002-500222/53.

XX
 PT New polypeptide comprising a death domain or death effector domain,
 PT useful for discovery of drugs that suppress infection, inflammation,
 PT allergy, sepsis, autoimmunity, allograft rejection and other diseases.

XX
 PS Example 7; Page 30; 99p; English.

XX
 CC The present invention provides novel death domain (DD) and death effector
 CC domain (DED) proteins and nucleic acids encoding them. The invention also
 CC provides death domain containing protein such as Chlamydia trachomatis
 CC death domain containing protein (CTDD) DD and neural growth factor
 CC receptor-interacting death domain (NIDD) DD. The invention is useful for
 CC identifying a binding agent (e.g. protein or drug) that binds a DD, DED
 CC or NB-ARC domain from DAP3, IRAK4, CTDD, DED4 or NIDD with a candidate
 CC binding agent and identifying an effective agent (e.g. protein or drug)
 CC that modulates the association of a DD, DED or NB-ARC domain with protein
 CC that binds the DD, DED or NB-ARC domain. The invention is also useful for
 CC modulating the level of cell process responses, such as apoptosis, cell adhesion,
 CC cell proliferation, cell stress responses, responses to microbial
 CC infection and B cell immunoglobulin class switching. DEDs, DEDs and NB-ARC
 CC domains and/or anti-DD, anti-DD or anti-NB-ARC domain antibodies are
 CC useful for discovery of drugs that suppress infection, autoimmunity,
 CC inflammation, allergy, allograft rejection, sepsis and other diseases.
 CC DD, DED or NB-ARC domain proteins are used to treat infection, allergy,
 CC autoimmunity, inflammation, allograft rejection, sepsis, keratinocyte
 CC hyperplasia, neoplasia, keloid, benign prostatic hypertrophy, fibrosis,
 CC inflammatory hyperplasia and smooth muscle cell proliferation in arteries
 CC following balloon angioplasty (restenosis). The invention is also used in
 CC antibody therapy and gene therapy. The present sequence is a primer used
 CC to amplify human DED4 cDNA

XX
 SQ Sequence 21 BP; 4 A; 8 C; 7 G; 2 T; 0 U; 0 Other;

Query Match 1.1%; Score 21; DB 6; Length 21;

Best Local Similarity 100.0%; Pred.No.45;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 CGCCGAGAGTCATCGAGCC 76
 ||||||||||||||||
 Db 1 CGCCGAGAGTCATCGAGCC 21

Search completed: March 23, 2005, 20:50:00
 Job time : 1099.22 secs

. 85 19 1.0 943 14 US-10-245-883-111 Sequence 111, App
86 19 1.0 943 14 US-10-237-535-111 Sequence 111, App
87 19 1.0 943 14 US-10-238-183-111 Sequence 111, App
88 19 1.0 943 14 US-10-238-283-111 Sequence 111, App
89 19 1.0 943 14 US-10-238-370-111 Sequence 111, App
90 19 1.0 943 14 US-10-245-055-111 Sequence 111, App
91 19 1.0 943 14 US-10-245-147-111 Sequence 111, App
92 19 1.0 943 14 US-10-245-730-111 Sequence 111, App
93 19 1.0 943 14 US-10-245-739-111 Sequence 111, App
94 19 1.0 943 14 US-10-246-210-111 Sequence 111, App
95 19 1.0 943 14 US-10-239-196-111 Sequence 111, App
96 19 1.0 943 14 US-10-243-024-111 Sequence 111, App
97 19 1.0 943 14 US-10-243-409-111 Sequence 111, App
98 19 1.0 943 14 US-10-245-621-111 Sequence 111, App
99 19 1.0 943 14 US-10-245-880-111 Sequence 111, App
100 19 1.0 943 14 US-10-245-033-111 Sequence 111, App

ALIGNMENTS

RESULT 1

US-10-296-539-4
; Sequence 4, Application US/10296539
; Publication No. US20030165933A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: DING, Li
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: BAUGHN, Marian R.
; TITLE OF INVENTION: REGULATORS OF APOPTOSIS
; FILE REFERENCE: PI-0307 PCT
; CURRENT APPLICATION NUMBER: US/10/296,539
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 60/209,407; 60/250,326
; PRIOR FILING DATE: 2000-06-01; 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030165933A1 3102521CB1
US-10-296-539-4

Query Match 60.0%; Score 1130; DB 16; Length 1230;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1180; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGCGTAATAATGAGAGGTCGAGAAAGATCCAAACAAAGTGGCTGCGCGCGCC 60
DB 4 AGCGTAATAATGAGAGGTCGAGAAAGATCCAAACAAAGTGGCTGCGCGCGCC 63
QY 61 AGGAGTCATCGAGCGCCAGATCTGGCGGGTCTGAGCTTGTTCGCGCTCCCTCCCG 120
DB 64 AGGAGTCATCGAGCGCCAGATCTGGCGGGTCTGAGCTTGTTCGCGCTCCCTCCCG 123
QY 121 GGAATGGCGCTATCCGGGTGACCCCGGCGCCGCTGCTGGAGAGAGATGATGCTGGAC 180
DB 124 GGAATGGCGCTATCCGGGTGACCCCGGCGCCGCTGCTGGAGAGAGATGATGCTGGAC 183
QY 181 TACTACGGAGATGCTGCTGCTTCAACCGATGTTGAGGTGGGCGGCAACTGACCGAG 240
DB 184 TACTACGGAGATGCTGCTGCTTCAACCGATGTTGAGGTGGGCGGCAACTGACCGAG 243
QY 241 TGGAGCTGGAGCTCTGGCGCTTCTGCTGATGAGGCTCTCGCGCGCGCGAGGCTTA 300

||||| 244 TCGAGCTGAGACTCTTGCGCTTTCGTGATGAGGCTCTCGCGCCGCCGAGGCTTA 303
QY 301 GCCCGGGCGCGAGGCGCTTAGAGCTCTGCTGAGCTGAGAGGCGCGGCGAGTGCAGC 360
DB 304 GCCCGGGCGCGAGGCGCTTAGAGCTCTGCTGAGCTGAGAGGCGCGGCGAGTGCAGC 363
QY 361 GAGAGCACTGCGCTGCTGAGGAGCACTCTGCGGTGCTGAGCCGCGAGCACTGCTG 420
DB 364 GAGAGCACTGCGCTGCTGAGGAGCACTCTGCGGTGCTGAGCCGCGAGCACTGCTG 423
QY 421 CCGCACTGCGCGGCAAGCGCGCGCGCGCACTGCTTCCAGAACCTATGCTATGCAAC 480
DB 424 CCGCACTGCGCGGCAAGCGCGCGCGCGCACTGCTTCCAGAACCTATGCTATGCAAC 483
QY 481 TCGAGCTTCAAAGAGCAGAGGCTAGCTGCGCGCGCGCTGCGGAGTCAAGCACTTCT 540
DB 484 TCGAGCTTCAAAGAGCAGAGGCTAGCTGCGCGCGCGCTGCGGAGTCAAGCACTTCT 543
QY 541 GCAAAATTCACAGAGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGAGCGCGG 600
DB 544 GCAAAATTCACAGAGGTCAGTGGAGACAGGCTCCCGCCCAACCAAGCGGAGCGCGG 603
QY 601 AGTGGGGCGCGCCCACTGCTGCTGAGACAGGCGCGGAGAGGCGCGCCAGCGGACCC 660
DB 604 AGTGGGGCGCGCCCACTGCTGCTGAGACAGGCGCGGAGAGGCGCGCCAGCGGACCC 663
QY 661 CAGCAGCACTGAGAGCGCGCGCGAGACCTTCTCTGAAAGGCAAGTGCCTGTGACATCCG 720
DB 664 CAGCAGCACTGAGAGCGCGCGCGAGACCTTCTCTGAAAGGCAAGTGCCTGTGACATCCG 723
QY 721 CTCGGGGTTCAGACAGATGACTGAGACAGTGGGCGCACTTGGAGAGAGGCGGTGACATC 780
DB 724 CTCGGGGTTCAGACAGATGACTGAGACAGTGGGCGCACTTGGAGAGAGGCGGTGACATC 783
QY 781 CCGCGGCGCCOAGCGCTGCGCGCGCGAGCTGAGAGTGTGTTGGGCAAGCCGAGATGCTG 840
DB 784 CCGCGGCGCCOAGCGCTGCGCGCGCGAGCTGAGAGTGTGTTGGGCAAGCCGAGATGCTG 843
QY 841 CGCTCAAGGAGCCTGGGCTCTGAGTGTGATGATCAATGATTCAGAGCTTCTCTATCTG 900
DB 844 CGCTCAAGGAGCCTGGGCTCTGAGTGTGATGATCAATGATTCAGAGCTTCTCTATCTG 903
QY 901 GACGCTTCTGAGGCGCACTGAGTGGCGGCTGCTGAGGCGGCGGCGGTGTC 960
DB 904 GACGCTTCTGAGGCGCACTGAGTGGCGGCTGCTGAGGCGGCGGCGGTGTC 963
QY 961 CTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGCTGCTGAGTGTG 1020
DB 964 CTGACTGAGGCGCTGCGAGAGGCTGTGGGCGGAGGCTGTTCGCTGCTGCTGAGTGTG 1023
QY 1021 GATGAGGCTGATGATGAGGCTGGCGGCGCGCGCTGCTGATGAGAGAGGAGGAGG 1080
DB 1024 GATGAGGCTGATGATGAGGCTGGCGGCGCGCGCTGCTGATGAGAGAGGAGGAGG 1083
QY 1081 CCGCGCGCGAGAGGCGCTCTGATCCAGACTGCGAGATTGATCCACTTCAAGTCT 1140
DB 1084 CCGCGCGCGAGAGGCGCTCTGATCCAGACTGCGAGATTGATCCACTTCAAGTCT 1143
QY 1141 CCGGCGCACTTCTCTGAGAGAGCAGCAATCTTACCCCT 1181
DB 1144 CCGGCGCACTTCTCTGAGAGAGCAGCAATCTTACCCCT 1184

RESULT 2

US-10-013-477-2
; Sequence 2, Application US/10013477
; Publication No. US20030049732A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Apoptosis Related Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT002PI
; CURRENT APPLICATION NUMBER: US/10/013,477

```
/ CURRENT FILING DATE: 2001-12-13
/ PRIOR APPLICATION NUMBER: 09/669,445
/ PRIOR FILING DATE: 2000-09-25
/ PRIOR APPLICATION NUMBER: PCT/US00/06642
/ PRIOR FILING DATE: 2000-03-15
/ PRIOR APPLICATION NUMBER: 60/126,018
/ PRIOR FILING DATE: 1999-03-24
/ PRIOR APPLICATION NUMBER: 60/139,638
/ PRIOR FILING DATE: 1999-06-17
/ PRIOR APPLICATION NUMBER: 60/149,449
/ PRIOR FILING DATE: 1999-08-18
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 2045
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-013-477-2
```

```
Query Match      58.7%; Score 1105; DB 14; Length 2045;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 26 GAAAGATCCAAACAAAGTGTGCTGCGGCGCTGCGCCAGAGATCATCGAGCCGAGAAATCTG 85
DB 23 GAAAGATCCAAACAAAGTGTGCTGCGGCGCTGCGCCAGAGATCATCGAGCCGAGAAATCTG 82
QY 86 GCCGGGTTCTGAGCTTGTTCCTGCTCCCTCCCGGGAATGCGGTATCCGGGTGACCC 145
DB 83 GCCGGGTTCTGAGCTTGTTCCTGCTCCCTCCCGGGAATGCGGTATCCGGGTGACCC 142
QY 146 CGGCCCCGTGCTGGAGAGATGAGTCTGAGCTTACCTTACGAGATCTTCCCTTAC 205
DB 143 CGGCCCCGTGCTGGAGAGATGAGTCTGAGCTTACCTTACGAGATCTTCCCTTAC 202
QY 206 GTATGTTGAGAGTGTGAGCGGCAACTGACCGAGTCTGAGTCTTCTGAGCTTTG 265
DB 203 GTATGTTGAGAGTGTGAGCGGCAACTGACCGAGTCTGAGTCTTCTGAGCTTTG 262
QY 266 TGTGTGATGAGGCTCTGCGCGCGCGGAGGCTTACGCGCGCGCGCGAGGCGGCTTACG 325
DB 263 TGTGTGATGAGGCTCTGCGCGCGCGGAGGCTTACGCGCGCGCGAGGCGGCTTACG 322
QY 326 TCTGTGTGAGCTGAGCGCGCGCGGAGGCTGCGCGGAGAGCAACTTCCGCTGCTGAGG 385
DB 323 TCTGTGTGAGCTGAGCGCGCGGAGGCTGCGCGGAGAGCAACTTCCGCTGCTGAGG 382
QY 386 AACTCTGCGAGTGTGAGCGCGCGGAGGCTGCGCGGAGAGCAACTTCCGCTGCTGAGG 445
DB 383 AACTCTGCGAGTGTGAGCGCGCGGAGGCTGCGCGGAGAGCAACTTCCGCTGCTGAGG 442
QY 446 GGCAGAGTGTCTCAGAAACGCTTATAGCTTATGAGCACTTCAAGAGAGAGAGAG 505
DB 443 GGCAGAGTGTCTCAGAAACGCTTATAGCTTATGAGCACTTCAAGAGAGAGAGAG 502
QY 506 GTAGCTGCGCTGCGCGCTGCGGAGTCAAGAGTCTTGAATTTCTCAGAGAGGTCAGTGG 565
DB 503 GTAGCTGCGCTGCGCGCTGCGGAGTCAAGAGTCTTGAATTTCTCAGAGAGGTCAGTGG 562
QY 566 AGACAGAGCTGCGCGCTGCGGAGTCAAGAGTCTTGAATTTCTCAGAGAGGTCAGTGG 625
DB 563 AGACAGAGCTGCGCGCTGCGGAGTCAAGAGTCTTGAATTTCTCAGAGAGGTCAGTGG 622
QY 626 CCAGAGCGGCGGAGAGAGGAGGCGCGGAGGCTGCGCGGAGAGAGTCAAGAGAGG 685
DB 623 CCAGAGCGGCGGAGAGAGGAGGCGCGGAGGCTGCGCGGAGAGAGTCAAGAGAGG 682
QY 686 CTTCCTCTGAAGAGCAAGTGAATCTGTGACATCCGGCTCCGGGTTTCAGAGAGATGAGG 745
DB 683 CTTCCTCTGAAGAGCAAGTGAATCTGTGACATCCGGCTCCGGGTTTCAGAGAGATGAGG 742
QY 746 AGCATGGGCGAGGCTTGAAGAGAGGAGTGTGAGTCCCGGCGGAGGCGGAGGCTGAGG 805
DB 746 AGCATGGGCGAGGCTTGAAGAGAGGAGTGTGAGTCCCGGCGGAGGCGGAGGCTGAGG 805
```

```
DB 743 AGCATGGGCGAGGCTTGAAGAGAGGAGTGTGAGTCCCGGCGGAGGCGGAGGCTGAGG 802
QY 806 AGCTGAGAGCTGTTTGGGCAAGGCCAAGGAGTGTGCGCTTCAAGGAGCTTGGCTTGTG 865
DB 803 AGCTGAGAGCTGTTTGGGCAAGGCCAAGGAGTGTGCGCTTCAAGGAGCTTGGCTTGTG 862
QY 866 TTTGTGACATCAAGTCTGAGAGCTTCCATCTGAGAGGCTTCTGAGGAGGAGTCACTGA 925
DB 863 TTTGTGACATCAAGTCTGAGAGCTTCCATCTGAGAGGCTTCTGAGGAGGAGTCACTGA 922
QY 926 GTGCGCGCTGCTGAGAGGCTTCCGCGGAGGCTTCTGAGAGGCTTCCGAGAGGCTG 985
DB 923 GTGCGCGCTGCTGAGAGGCTTCCGCGGAGGCTTCTGAGAGGCTTCCGAGAGGCTG 982
QY 986 TGGGCGGAGGCTGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1045
DB 983 TGGGCGGAGGCTGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1042
QY 1046 GGGCGCGCTGCTGAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1105
DB 1043 GGGCGCGCTGTTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 1102
QY 1106 CCAGAGCTGAGAGATGATCCCACTTCAAGTCTCCGAGGCACTTCTTCTGAGAGAG 1165
DB 1103 CCAGAGCTGAGAGATGATCCCACTTCAAGTCTCCGAGGCACTTCTTCTGAGAGAG 1162
QY 1166 GACCATCTTACCCCT 1181
DB 1163 GACCATCTTACCCCT 1178
```

RESULT 3

```
US-10-106-698-349
/ Sequence 349, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptid
/ FILE REFERENCE: PA005F1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: Patentln Ver. 3.0
/ SEQ ID NO 349
/ LENGTH: 2045
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-106-698-349
```

```
Query Match      58.7%; Score 1105; DB 15; Length 2045;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1155; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 26 GAAAGATCCAAACAAAGTGTGCTGCGGCGCTGCGCCAGAGATCATCGAGCCGAGAAATCTG 85
DB 23 GAAAGATCCAAACAAAGTGTGCTGCGGCGCTGCGCCAGAGATCATCGAGCCGAGAAATCTG 82
QY 86 GCCGGGTTCTGAGCTTGTTCCTGCTCCCTCCCGGGAATGCGGTATCCGGGTGACCC 145
DB 83 GCCGGGTTCTGAGCTTGTTCCTGCTCCCTCCCGGGAATGCGGTATCCGGGTGACCC 142
QY 146 CGGCCCCGTGCTGGAGAGAGATGAGTCTTGAATTTCTCAGAGAGGTCAGTGG 205
DB 143 CGGCCCCGTGCTGGAGAGAGATGAGTCTTGAATTTCTCAGAGAGGTCAGTGG 202
QY 206 GTATGTTGAGAGTGTGAGCGGCAACTGACCGAGTCTGAGTCTTCTGAGCTTTG 265
DB 206 GTATGTTGAGAGTGTGAGCGGCAACTGACCGAGTCTGAGTCTTCTGAGCTTTG 265
```

Db 203 GATATGTCGAGGTGATGAGGCGGCAACTGACGAGTGCAGTCTGAGCTCTGCGCTTC 262
 QY 266 TGCTGATGAGGCTCTGAGGCGCGCGAGGCTTAGCCGCGGCGCGAGCGGCTGAGG 325
 Db 263 TGCTGATGAGGCTCTGAGGCGCGCGAGGCTTAGCCGCGGCGCGAGCGGCTGAGG 322
 QY 326 TCCTGCTGAGCTGAGGCGCGCGAGGCTGAGGCGAGCACTGCGGCTGCTGAGG 385
 Db 323 TCCTGCTGAGCTGAGGCGCGCGAGGCTGAGGCGAGCACTGCGGCTGCTGAGG 382
 QY 386 AACTCTGAGGCTGAGGCGCGCGAGGCTGAGGCGAGCACTGCGGCTGCTGAGG 445
 Db 383 AACTCTGAGGCTGAGGCGCGCGAGGCTGAGGCGAGCACTGCGGCTGCTGAGG 442
 QY 446 GGCAGTGTCTCAGAGACGTATGATGAGGCACTTCAAGCTTCTCAAGAGAGAGAG 505
 Db 443 GGCAGTGTCTCAGAGACGTATGATGAGGCACTTCAAGCTTCTCAAGAGAGAGAG 502
 QY 506 GTAGCTGCGCTGCGCTGCGAGTCAAGCAAGTTCTGCAATTCTCAGCAGAGGTCA 565
 Db 503 GTAGCTGCGCTGCGCTGCGAGTCAAGCAAGTTCTGCAATTCTCAGCAGAGGTCA 562
 QY 566 AGACAGGCTTCCCCCAACCAACCGGAGGCGGAGTGGGGCGCGCCAGAGTGTGTG 625
 Db 563 AGACAGGCTTCCCCCAACCAACCGGAGGCGGAGTGGGGCGCGCCAGAGTGTGTG 622
 QY 626 CCAGACGCGCGGAGAGGCGGCGCGAGCGGCGAGCGAGAGTCAAGAGCGCGAGAG 685
 Db 623 CCAGACGCGCGGAGAGGCGGCGCGAGCGGCGAGCGAGAGTCAAGAGCGCGAGAG 682
 QY 686 CTTCCTCTGAGGCAAGTGAACCTGTGACATCCGGCTCCGGGTTCTGAGCAGAGTACTG 745
 Db 683 CTTCCTCTGAGGCAAGTGAACCTGTGACATCCGGCTCCGGGTTCTGAGCAGAGTACTG 742
 QY 746 AGATAGGCGGAGCTTGGAGAGAGGCGGTGGACATCCCGGCGCGCGCGAGCGGCG 805
 Db 743 AGATAGGCGGAGCTTGGAGAGAGGCGGTGGACATCCCGGCGCGCGCGAGCGGCG 802
 QY 806 AGCTGAGCGTGTGGAGAGAGGCGGAGGCGAGTGTGCGCTCAAGAGAGCTTGGGCTGTG 865
 Db 803 AGCTGAGCGTGTGGAGAGAGGCGGAGGCGAGTGTGCGCTCAAGAGAGCTTGGGCTGTG 862
 QY 866 TTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGGCTTCTGGGCGAGTACTGA 925
 Db 863 TTGTGACATCAAGTTCTCAGAGCTCTCTATCTGAGCGGCTTCTGGGCGAGTACTGA 922
 QY 926 GTGCGCGCTGCTGAGGCGGCGGCGGCTGCTGAGCTGAGGCGGCTGCGAGAGCTG 985
 Db 923 GTGCGCGCTGCTGAGGCGGCGGCGGCTGCTGAGCTGAGGCGGCTGCGAGAGCTG 982
 QY 986 TGAGCGGAGAGGCTGTTGCGCTCTGATGAGTGTGAGTGAAGCTGATGAGGCTGAGG 1045
 Db 983 TGAGCGGAGAGGCTGTTGCGCTCTGATGAGTGTGAGTGAAGCTGATGAGGCTGAGG 1042
 QY 1046 GGGCGCGCTGTTGCTGATGAGAGAGAGAGGCGGCGCGCGAGAGAGGCTCTGTAT 1105
 Db 1043 GGGCGCGCTGTTGCTGATGAGAGAGAGAGAGGCGGCGCGCGAGAGAGGCTCTGTAT 1102
 QY 1106 CCGAGACTGCGAGAGATTTGATCCCACTCCAAAGTCTCGGGGCGAGCTTCTCTGGAGAG 1165
 Db 1103 CCGAGACTGCGAGAGATTTGATCCCACTCCAAAGTCTCGGGGCGAGCTTCTCTGGAGAG 1162
 QY 1166 GACCATCTTACCCCT 1181
 Db 1163 GACCATCTTACCCCT 1178

RESULT 4
 US-10-001-254-17
 ; Sequence 17, Application US/10001254
 ; Publication No. US20030049702A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.

; APPLICANT: Godzik, Adam
 ; APPLICANT: Pawlowski, Krzysztof
 ; APPLICANT: Fiorentino, Irodana
 ; APPLICANT: Lee, Sung Hyung
 ; APPLICANT: Roth, Wilfred
 ; APPLICANT: Steiner-Liwen, Frank
 ; TITLE OF INVENTION: No. US20030049702A1 Death Domain Proteins
 ; FILE REFERENCE: P-Ld 5037
 ; CURRENT APPLICATION NUMBER: US/10/001,254
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: 60/301,889
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/715,893
 ; NUMBER OF SEQ ID NOS: 62
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 1924
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (91)..(1044)
 ; NAME/KEY: misc_Feature
 ; LOCATION: (1)..(1900)
 ; OTHER INFORMATION: n = A,T,C or G
 US-10-001-254-17

Query Match 57.2%; Score 1078; DB 14; Length 1924;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1128; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 53 CGTGGCCGAGAGTCAATCGAGCGGAGAGTCTGAGCGGCTTCTGAGCTTCTGCGCTCC 112
 Db 20 CGTGGCCGAGAGTCAATCGAGCGGAGAGTCTGAGCGGCTTCTGAGCTTCTGCGCTCC 79
 QY 113 CTCGCCGCGGAATGCGGCTATCCGGGTGAGCCCGGCGCGCTGCTGAGAGAGTGA 172
 Db 80 CTCGCCGCGGAATGCGGCTATCCGGGTGAGCCCGGCGCGCTGCTGAGAGAGTGA 139
 QY 173 GCTTGACATCAAGGAGAGTGTGTGCTTCAACCTATGTTCCAGTGTGTGGCGGCGAAC 232
 Db 140 GCTTGACATCAAGGAGAGTGTGTGCTTCAACCTATGTTCCAGTGTGTGGCGGCGAAC 199
 QY 233 TGACGAGTGCAGCTGAGAGCTCTGCGCTTCTGCTGATGAGGCTCTGCGCGCGCG 292
 Db 200 TGACGAGTGCAGCTGAGAGCTCTGCGCTTCTGCTGATGAGGCTCTGCGCGCGCG 259
 QY 293 GAGGCTTACCCGCGCGCGGCGGCTGAGAGCTCTGCTGAGAGCTGAGCGCGCGGCG 352
 Db 260 GAGGCTTACCCGCGCGCGGCGGCTGAGAGCTCTGCTGAGAGCTGAGCGCGCGGCG 319
 QY 353 AGTGGCGGAGAGCAACTGCGGCTGCTGAGGCACTCTGCGGCTGCTGCGCGCGGCG 412
 Db 320 AGTGGCGGAGAGCAACTGCGGCTGCTGAGGCACTCTGCGGCTGCTGCGCGCGGCG 379
 QY 413 ACCTGTCGCGCACTGCGGCGGAGAGCGGCGCGGCGGCACTGCTCTCAGAGCTTATAGCT 472
 Db 380 ACCTGTCGCGCACTGCGGCGGAGAGCGGCGCGGCGGCACTGCTCTCAGAGCTTATAGCT 439
 QY 473 ATGGCACTTCACGCTTCTCAAGAGAGAGAGAGTGTGCGGTGCGGCTGCGGAGTCA 532
 Db 440 ATGGCACTTCACGCTTCTCAAGAGAGAGAGAGTGTGCGGTGCGGCTGCGGAGTCA 499
 QY 533 GCACTTCTGCAATTCTCAGCAGGCTCAATGGAGAGCAGGCTTCTGCGGCGGCGG 592
 Db 500 GCACTTCTGCAATTCTCAGCAGGCTCAATGGAGAGCAGGCTTCTGCGGCGGCGG 559
 QY 593 AGCGCGGAGTGGGGCGGCGCGGAGTGTGCGGAGAGCGGCGGAGAGGAGGCGGCGAG 652
 Db 560 AGCGCGGAGTGGGGCGGCGCGGAGTGTGCGGAGAGCGGCGGAGAGGAGGCGGCGAG 619
 QY 653 CCGCACCCGAGAGAGTGAAGCCCGGCGGAGCTTCTCTGAGAGCAAGTGAAGTGA 712

Db 620 CCGCACCCGAGCGAGTGAAGCCGCGCAGACCTTCTCTGAAGGAAAGTACCTGTG 679
Qy 713 ACATCCGAGCTCCGGGTTTCAGACAGATCTGCCAGATGAGGCAAGCTTGAAGAGGCG 772
Db 680 ACATCCGAGCTCCGGGTTTCAGACAGATCTGCCAGATGAGGCAAGCTTGAAGAGGCG 739
Qy 773 TGGCATCCCGGCGGCGCCGAGGCGCTGGCGGCGAGCTGAGAGTGTGTTGGGCAAGCCACG 832
Db 740 TGGCATCCCGGCGGCGCCGAGGCGCTGGCGGCGAGCTGAGAGTGTGTTGGGCAAGCCACG 799
Qy 833 CAGTGTGCGCTCAAGGAGCCTGAGCTGTGTGTTTGTGACATCAAGTTCTCAGAGCTCT 892
Db 800 CAGTGTGCGCTCAAGGAGCCTGAGGCTGTGTGTTTGTGACATCAAGTTCTCAGAGCTCT 859
Qy 893 CCTATCTGAGCGCTTCTGAGGCGACTACTGATGAGCGCTCTGCTGCAAGCCCTGCGG 952
Db 860 CCTATCTGAGCGCTTCTGAGGCGACTACTGATGAGCGCTCTGCTGCAAGCCCTGCGG 919
Qy 953 GCGTGTCTGACTGAGGCGCTGAGAGGCTGTGGGCGGAGGCGCTGTGCGCTGTG 1012
Db 920 GCGTGTCTGACTGAGGCGCTGAGAGGCTGTGGGCGGAGGCGCTGTGCGCTGTG 979
Qy 1013 TCAGTGTGATGAGGCTGACTATGAGGCTGGCGGCGCGCTGTGCTGATGAGAGG 1072
Db 980 TCAGTGTGATGAGGCTGACTATGAGGCTGGCGGCGCGCTGTGCTGATGAGAGG 1039
Qy 1073 AAGGGGGCGGCGCGCGAGAGGCTCTGATTCAGAGATGAGAGATGATCCACT 1132
Db 1040 AAGGGGGCGGCGCGCGAGAGGCTCTGATTCAGAGATGAGAGATGATCCACT 1099
Qy 1133 CCAAGTCTCCGGGCGACCTTCTCTGGGAGAGAGCAACCATCTCAACCT 1181
Db 1100 CCAAGTCTCCGGGCGACCTTCTCTGGGAGAGAGCAACCATCTCAACCT 1148

RESULT 5
US-09-925-302-315
; Sequence 315, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 315
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-315

Query Match 52.3%; Score 984; DB 9; Length 2044;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1154; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 26 GAAAGATCCAAACCAAGTGGCTGCGGCGCTGCGCCAGAGATCATCGAGCGCCAGAACTGTG 85
Db 23 GAAAGATCCAAACCAAGTGGCTGCGGCGCTGCGCCAGAGATCATCGAGCGCCAGAACTGTG 82
Qy 86 GCGGAGTTCTGAGCTTGTTCGCGCTCCCTCCCGGGAATGCGCTATCCGGGTGACGCC 145
Db 83 GCGGAGTTCTGAGCTTGTTCGCGCTCCCTCCCGGGAATGCGCTATCCGGGTGACGCC 142
Qy 146 CCGCCCGGTGCTGGAGAGAGATGAGTCTGTGACTATCAAGGATGCTGTGCTTCACC 205
Db 143 CCGCCCGGTGCTGGAGAGAGATGAGTCTGTGACTATCAAGGATGCTGTGCTTCACC 202

Qy 206 GTATGTTCCAGAGTGTGGGCGGCGCACTGACCGAGTGCAGCTGAGACTTCGTGGCCTTTC 265
Db 203 GTATGTTCCAGAGTGTGGGCGGCGCACTGACCGAGTGCAGCTGAGACTTCGTGGCCTTTC 262
Qy 266 TGTGTGATGAGGCTCTCTGGGCGCGCGAGGCTTTAAGCCCGGAGCCCGAGCGGCTAGAGC 325
Db 263 TGTGTGATGAGGCTCTCTGGGCGCGCGAGGCTTTAAGCCCGGAGCCCGAGCGGCTAGAGC 322
Qy 326 TCTGTGTGAGCTTGAAGCGCGCGGCGAGTGGCGGCGAGACAACTTCCGCTCTGTGGGC 385
Db 323 TCTGTGTGAGCTTGAAGCGCGCGGCGAGTGGCGGCGAGACAACTTCCGCTCTGTGGGC 382
Qy 386 AACTCTGCGCGGTGCTGGCGCGCGGCGAGACCTGTGCGCGCACTGGCGCGGAGCGCGGC 445
Db 383 AACTCTGCGCGGTGCTGGCGCGCGGCGAGACCTGTGCGCGCACTGGCGCGGAGCGCGGC 442
Qy 446 GGCAGATGCTCCAGAACGCTATAGCTATGAGCACTTCAGACTTTCAAAGAGACAGAG 505
Db 443 GGCAGATGCTCCAGAACGCTATAGCTATGAGCACTTCAGACTTTCAAAGAGACAGAG 502
Qy 506 GTAGTGTGCGTCCGCTCGGAGTCAAGCACTTCTGCAAAATTCTCAGCAGGCTCAGTGG 565
Db 503 GTAGTGTGCGTCCGCTCGGAGTCAAGCACTTCTGCAAAATTCTCAGCAGGCTCAGTGG 562
Qy 566 AGACAGGCTCCCGCCCAACCAAGCGGCGAGCGGAGTGGGCGCGCGCCCAAGTGTGTG 625
Db 563 AGACAGGCTCCCGCCCAACCAAGCGGCGAGCGGAGTGGGCGCGCGCCCAAGTGTGTG 622
Qy 626 CCAGAGCGCGCGAGAGAGGCGCCCAAGCGCACCCCAAGCAGTCAAGAGCCCGCCAGAC 685
Db 623 CCAGAGCGCGCGAGAGAGGCGCCCAAGCGCACCCCAAGCAGTCAAGAGCCCGCCAGAC 682
Qy 686 CTTCTCTGAAAGCAAGTACCTGTGACATCCGCTCCGGGTTTCAAGAGAGTACTGCG 745
Db 683 CTTCTCTGAAAGCAAGTACCTGTGACATCCGCTCCGGGTTTCAAGAGAGTACTGCG 742
Qy 746 AGCATGGGCGAGGCTTGGAGAGAGGCGTGGCAATCCCGGCGCGCCCGAGCGCTGGCGGC 805
Db 743 AGCATGGGCGAGGCTTGGAGAGAGGCGTGGCAATCCCGGCGCGCCCGAGCGCTGGCGGC 802
Qy 806 AGCTGAGCGTGTGGGCGAGGCGCACCGCAGTGTGCGCTCAAGAGGACTTGGGCTTGTGG 865
Db 803 AGCTGAGCGTGTGGGCGAGGCGCACCGCAGTGTGCGCTCAAGAGGACTTGGGCTTGTGG 862
Qy 866 TTTGTGACATCAAGATTCTCAGAGCTTCTTATGTGACGCTTCTGAGGCGACTTACTGA 925
Db 863 TTTGTGACATCAAGATTCTCAGAGCTTCTTATGTGACGCTTCTGAGGCGACTTACTGA 922
Qy 926 GTGGCGCGCTGTGCGAGCGCTTGGCGGCGTGTCTGTGACTGAGGCGCTTGGAGAGGCTG 985
Db 923 GTGGCGCGCTGTGCGA-GCCCTGCGGCGCTGTCTGTGACTGAGGCGCTTGGAGAGGCTG 981
Qy 986 TGGGCGGAGGCTGTTTGGCTGCTGCTGATGAGTGTGATGAGGCTGATATGAGGCTTGGC 1045
Db 982 TGGGCGGAGGCTGTTTGGCTGCTGATGAGTGTGATGAGGCTGATATGAGGCTTGGC 1041
Qy 1046 GGGCGCGCTGTTTGTGATGAGAGAGAGGAGGCGCGCGCCGACAGAGGCTTCTGTAT 1105
Db 1042 GGGCGCGCTGTTTGTGATGAGAGAGAGGAGGCGCGCGCCGACAGAGGCTTCTGTAT 1101
Qy 1106 CAGAGACTGGCAGAGATTGATCCCACTTCAAGTCTTCCGGGCGACCTTCTCTGGAGAGAC 1165
Db 1102 CAGAGACTGGCAGAGATTGATCCCACTTCAAGTCTTCCGGGCGACCTTCTCTGGAGAGAC 1161
Qy 1166 GACCATCTCAACCTCT 1181
Db 1162 GACCATCTCAACCTCT 1177

RESULT 6
US-09-925-302-315
; Sequence 315, Application US/09925302

```

; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 315
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-315

```

```

Query Match          52.3%; Score 984; DB 10; Length 2044;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1154; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

```

QY 26 GAAAGATCCAAAACAAGTGGCTCGGCGCTGCCCAAGAGTATCGAGCGCCAGAAATCTG 85
DB 23 GAAAGATCCAAAACAAGTGGCTCGGCGCTGCCCAAGAGTATCGAGCGCCAGAAATCTG 82
QY 86 GCGGGGTTCTGACCTGTTCCGCTCCCTCCCGGGGAATGGGGCTATCCGGGTGACCC 145
DB 83 GCGGGGTTCTGACCTGTTCCGCTCCCTCCCGGGGAATGGGGCTATCCGGGTGACCC 142
QY 146 CCGCCCCGCTGCTGGAGAGAGATGATGCTGTGACTACTACGAGATGCTGTCCCTTAC 205
DB 143 CCGCCCCGCTGCTGGAGAGAGATGATGCTGTGACTACTACGAGATGCTGTCCCTTAC 202
QY 206 GTATGTTCAAGTGTGTGGCGGCACTGACCGAGTGCAGTGTGAGCTTCTGGCTTTC 285
DB 203 GTATGTTCAAGTGTGTGGCGGCACTGACCGAGTGCAGTGTGAGCTTCTGGCTTTC 282
QY 266 TGTGTATAGAGGTCTCTGGCGCGCGCGGAGCTTACCGCGGGCCCAAGCGCTGAGC 325
DB 263 TGTGTATAGAGGTCTCTGGCGCGCGCGGAGCTTACCGCGGGCCCAAGCGCTGAGC 322
QY 326 TCTCTGTGAGCTGAGCGCGCGGAGTGCAGAGCACTGCGGCTGCTGAGC 385
DB 323 TCTCTGTGAGCTGAGCGCGCGGAGTGCAGAGCACTGCGGCTGCTGAGC 382
QY 386 AACTCTGCGCTGCTGCTGCGCGCACTGCTGCGCACTTGGCGCGCAACGGCGCC 445
DB 383 AACTCTGCGCTGCTGCTGCGCGCACTGCTGCGCACTTGGCGCGCGCAACGGCGCC 442
QY 446 GGCAGTGTCTCCAGAACGCTATAGTATGAGCACTTCCAGCTTTCAGAGAGAGAG 505
DB 443 GGCAGTGTCTCCAGAACGCTATAGTATGAGCACTTCCAGCTTTCAGAGAGAGAG 502
QY 506 GTAGTGTCCGTGCGCGTGCAGTCAAGCACTTCTGCAATTCTCAGCAGGCTCAGTGG 565
DB 503 GTAGTGTCCGTGCGCGTGCAGTCAAGCACTTCTGCAATTCTCAGCAGGCTCAGTGG 562
QY 566 AGACAGGCTTCCCCCAACCAAGCGGAGCGGAGTGTGGGCGCGGCCAGTGTGGTG 625
DB 563 AGACAGGCTTCCCCCAACCAAGCGGAGCGGAGTGTGGGCGCGGCCAGTGTGGTG 622
QY 626 CCAAGAGGCGGAGAGAGGCGCGCCAGCGCACCCCAAGCAGCAGTCAAGCGCGCCAG 685
DB 623 CCAAGAGGCGGAGAGAGGCGCGCCAGCGCACCCCAAGCAGCAGTCAAGCGCGCCAG 682
QY 686 CTTCCTCTGAAGCAAGTGAAGTGTGACATCCGGCTCCGGGTTCAGCAGAGTACTGCG 745
DB 683 CTTCCTCTGAAGCAAGTGAAGTGTGACATCCGGCTCCGGGTTCAGCAGAGTACTGCG 742
QY 746 AGCATGGGCCAGCTTTGAGCAGAGGCGTGGCATCCCGGCGCCCAAGCGCTGGCGGC 805

```

```

DB 743 AGCATGGGCCAGCTTTGAGCAGAGGCGTGGCATCCCGGCGCCCAAGCGCTGGCGGC 802
QY 806 AGCTGACGCTGTTTGGGAGAGCCACCGCAGTGTGCTCAAGAGCACTGGGCTGTGG 865
DB 803 AGCTGACGCTGTTTGGGAGAGCCACCGCAGTGTGCTCAAGAGCACTGGGCTGTGG 862
QY 866 TTTGTGACATCAAGTTCAGAGCTTCTTATCTGAGCGCTTCTGGGCGACTACTGA 925
DB 863 TTTGTGACATCAAGTTCAGAGCTTCTTATCTGAGCGCTTCTGGGCGACTACTGA 922
QY 926 GTGGCGCTTGTGACAGCGCTTCCGGGGGCTGTCTTCTGACTAGAGCGCTCGAGAG 985
DB 923 GTGGCGCTTGTGACAGCGCTTCCGGGGGCTGTCTTCTGACTAGAGCGCTCGAGAG 981
QY 986 TGGGCGGAGAGCTGTTCCGCTGCTGTGTGATGTGATAGAGCTGATGAGCTGAGC 1045
DB 982 TGGGCGGAGAGCTGTTCCGCTGCTGTGTGATGTGATAGAGCTGATGAGCTGAGC 1041
QY 1046 GCGCGCGCTGTTGCTGATGAGAGAGAAAGGGGCGGCGCCGACAGAGGCTCTGAT 1105
DB 1042 GCGCGCGCTGTTGCTGATGAGAGAGAAAGGGGCGGCGCCGACAGAGGCTCTGAT 1101
QY 1106 CCAAGACTGGCAGATGATCCCACTCCAGTCTCCGGGCCACCTTCTCTGGAGAGC 1165
DB 1102 CCAAGACTGGCAGATGATCCCACTCCAGTCTCCGGGCCACCTTCTCTGGAGAGC 1161
QY 1166 GACCATCTACCCCT 1181
DB 1162 GACCATCTACCCCT 1177

```

```

RESULT 7
US-10-641-182
; Sequence 182, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
;             Susan G. Stuart
;             Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
;             GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37 071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 845-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 182:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: FIBRONOT01
 CLONE: 053532
 SEQUENCE DESCRIPTION: SEQ ID NO: 182 :
 US-10-641-643-182

Query Match 37.5%; Score 707; DB 17; Length 847;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1177 CCCCTTGACAGCCCTCCACAGAGATGGGCTCTGAGGCTTAAACCATTTCCAGCTGAG 1236
 DB 141 CCCCTTGACAGCCCTCCACAGAGATGGGCTCTGAGGCTTAAACCATTTCCAGCTGAG 200
 QY 1237 TTTCCTTCCAGACTCTCTCTACCCCGAGGTGCCCCCTTAAAGCTCCGAGGCGGGGC 1296
 DB 201 TTTCCTTCCAGACTCTCTCTACCCCGAGGTGCCCCCTTAAAGCTCCGAGGCGGGGC 260
 QY 1297 TGGGCGCTGATCTCAGAAAGGAGGAGGAGCACTACACCAAAAGGCGCCCTGAC 1356
 DB 261 TGGGCGCTGATCTCAGAAAGGAGGAGGAGCACTACACCAAAAGGCGCCCTGAC 320
 QY 1357 ATTGATCTCTGATCTTGGGCTGTCTGCACTGTCAAGTGCAACAACCTGCTCATGCTC 1416
 DB 321 ATTGATCTCTGATCTTGGGCTGTCTGCACTGTCAAGTGCAACAACCTGCTCATGCTC 380
 QY 1417 ACACGCGCCCTGCTAGATCTTCCCTGGGCTCTGCGCTGCTTCCGAGACAC 1476
 DB 381 ACACGCGCCCTGCTAGATCTTCCCTGGGCTCTGCGCTGCTTCCGAGACAC 440
 QY 1477 TTCTTTGGGCTTAAGGGCTTCTCTCTCAGACCTTATTTGACCAACAACCACTGGGCT 1536
 DB 441 TTCTTTGGGCTTAAGGGCTTCTCTCTCAGACCTTATTTGACCAACAACCACTGGGCT 500
 QY 1537 TCAGCCACATCATGTGGGCACTGAGAGTGGGTCGACATGGGCGCTGCTCACTTGGCCAC 1596
 DB 501 TCAGCCACATCATGTGGGCACTGAGAGTGGGTCGACATGGGCGCTGCTCACTTGGCCAC 560
 QY 1597 ACATCTCCAGCCAGCCAGGCGCCCTGCGCAGCTTCAATTTACAGACTGACTCTCTCACC 1656
 DB 561 ACATCTCCAGCCAGCCAGGCGCCCTGCGCAGCTTCAATTTACAGACTGACTCTCTCACC 620
 QY 1657 TTCCCCCTGCTGTCCAGACTGAACATAGACTTGGCACTTGGATGTCACTTGAAGTCA 1716
 DB 621 TTCCCCCTGCTGTCCAGACTGAACATAGACTTGGCACTTGGATGTCACTTGAAGTCA 680
 QY 1717 CATGGAGGTGTTATGGAGCATCATACCAAGGCTTATGTCATGGGCGCAAAACCA 1776
 DB 681 CATGGAGGTGTTATGGAGCATCATACCAAGGCTTATGTCATGGGCGCAAAACCA 740
 QY 1777 GTAAACAGCCACCTTCTTGGAAAGGAATCAAAAGGCTTTGGGGGTGATGAAAAAGACT 1836
 DB 741 GTAAACAGCCACCTTCTTGGAAAGGAATCAAAAGGCTTTGGGGGTGATGAAAAAGACT 800
 QY 1837 TTTTAAACAATGATACCAATTAACTGCGCTGAAAAAGGCGCATAGTGGG 1883
 DB 801 TTTTAAACAATGATACCAATTAACTGCGCTGAAAAAGGCGCATAGTGGG 847

RESULT 8

US-09-799-777-103
 ; Sequence 103, Application US/09799777
 ; Patent No. US20020091244A1

GENERAL INFORMATION:

APPLICANT: Lal, Preeti
 Hillman, Jennifer L.

Corley, Neil C.
 Guejler, Karl J.

Baugh, Mariah
 Sather, Susan

Shah, Purvi

TITLE OF INVENTION: HUMAN SIGNAL PEPTIDE-CONTAINING PROTEINS

NUMBER OF SEQUENCES: 154
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/799,777

FILING DATE: 06-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/002,485

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: BILLINGS, LUCY J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0459 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 103:

SEQUENCE CHARACTERISTICS:

LENGTH: 1554 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: UTRSNOT06

CLONE: 1638407

SEQUENCE DESCRIPTION: SEQ ID NO: 103 :

US-09-799-777-103

Query Match 37.5%; Score 707; DB 9; Length 1554;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1177 CCCCTTGACAGCCCTCCACAGAGATGGGCTCTGAGGCTTAAACCATTTCCAGCTGAG 1236
 DB 838 CCCCTTGACAGCCCTCCACAGAGATGGGCTCTGAGGCTTAAACCATTTCCAGCTGAG 897
 QY 1237 TTTCCTTCCAGACTCTCTCTACCCCGAGGTGCCCCCTTAAAGCTCCGAGGCGGGGC 1296
 DB 898 TTTCCTTCCAGACTCTCTCTACCCCGAGGTGCCCCCTTAAAGCTCCGAGGCGGGGC 957
 QY 1297 TGGGCGCTGATCTCAGAAAGGAGGAGGAGCACTACACCAAAAGGCGCCCTGAC 1356
 DB 958 TGGGCGCTGATCTCAGAAAGGAGGAGGAGCACTACACCAAAAGGCGCCCTGAC 1017
 QY 1357 ATTGATCTCTGATCTTGGGCTGTCTGCACTGTCAAGTGCAACAACCTGCTCATGCTC 1416
 DB 1018 ATTGATCTCTGATCTTGGGCTGTCTGCACTGTCAAGTGCAACAACCTGCTCATGCTC 1077
 QY 1417 ACACGCGCCCTGCTAGATCTTCCCTGGGCTCTGCGCTGCTTCCGAGACAC 1476
 DB 1078 ACACGCGCCCTGCTAGATCTTCCCTGGGCTCTGCGCTGCTTCCGAGACAC 1137
 QY 1477 TTCTTTGGGCTTAAGGGCTTCTCTCTCAGACCTTATTTGACCAACAACCACTGGGCT 1536
 DB 1138 TTCTTTGGGCTTAAGGGCTTCTCTCTCAGACCTTATTTGACCAACAACCACTGGGCT 1197
 QY 1537 TCAGCCACATCATGTGGGCACTGAGAGTGGGTCGACATGGGCGCTGCTCACTTGGCCAC 1596
 DB 1198 TCAGCCACATCATGTGGGCACTGAGAGTGGGTCGACATGGGCGCTGCTCACTTGGCCAC 1257
 QY 1597 ACATCTCCAGCCAGCCAGGCGCCCTGCGCAGCTTCAATTTACAGACTGACTCTCTCACC 1656


```
Db 376 GGGCCAGCCTTGGAGGAGGCGTGGCATCCGGGGGCGCCAGGGGCTGGGGCGGAGCTG 435
Qy 811 GAGCTGTTGGGCAAGCCACCGCACTGCTGCTCAAGGGAAGCTGGGCTCTGTGTTGT 870
Db 436 GAGGTGTTGGGCAAGCCACCGCACTGCTGCTCAAGGGAAGCTGGGCTCTGTGTTGT 495
Qy 871 GACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCCCTTGGGGGCGACTACCTGAGTGGC 930
Db 496 GACATCAAGTTCTCAGAGCTCTCTATCTGAGCGCCCTTGGGGGCGACTACCTGAGTGGC 555
Qy 931 GCCCTGTCGAGGCGCCCTGCGGGGCTGTTCTTGAAGGCGCCCTGCGAGAGGCTGTGGC 990
Db 556 GCCCTGTCGAGGCGCCCTGCGGGGCTGTTCTTGAAGGCGCCCTGCGAGAGGCTGTGGC 615
Qy 991 CGGAGAGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
Db 616 CGGAGAGCTGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 675
Qy 1051 CGCCTGTTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1110
Db 676 CGCCTGTTGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 735
Qy 1111 ACTGGAGAGATTGATCCCACTCCAAAGTCTCCGGGCGCACTTCTCTGGAGAGAGCA 1170
Db 736 ACTGGAGAGATTGATCCCACTCCAAAGTCTCCGGGCGCACTTCTCTGGAGAGAGCA 795
Qy 1171 TCTCTACCCCT 1181
Db 796 TCTCTACCCCT 806

RESULT 11
US-09-796-692-2721
; Sequence 2721, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2721
; LENGTH: 523
```

```
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (15)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (502)
OTHER INFORMATION: n=A,T,C or G
US-09-796-692-2721

Query Match 23.1%; Score 435; DB 9; Length 523;
Best Local Similarity 99.8%; Pred. No. 1.4e-214;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1247 AGACTCTCTTACCCCAAGGTGTGCCCCCTTACCTTCGGAGGCGGGGGCTGGGCTGTGA 1306
Db 16 AGACTCTCTTACCCCAAGGTGTGCCCCCTTACCTTCGGAGGCGGGGGCTGGGCTGTGA 75
Qy 1307 TCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1366
Db 76 TCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 135
Qy 1367 TGATCTTGGGCTGTGCTGCACTGTGCAAGGTGCAACAAGTGTGATGCTGCACTGGCCC 1426
Db 136 TGATCTTGGGCTGTGCTGCACTGTGCAAGGTGCAACAAGTGTGATGCTGCACTGGCCC 135
Qy 1427 TGCTGAGATTTTCCCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1486
Db 196 TGCTGAGATTTTCCCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 255
Qy 1487 TAAGGCTTCTCTCAGGAGCTCTAATTGACCAACCAACCAACCAACCAACCAACCAACCA 1546
Db 256 TAAGGCTTCTCTCAGGAGCTCTAATTGACCAACCAACCAACCAACCAACCAACCAACCA 315
Qy 1547 CAGTGGGCACTGAGGTGGGGTGCACATGGGGCTGCTCACTTGGCCACACATCTTCAG 1606
Db 316 CAGTGGGCACTGAGGTGGGGTGCACATGGGGCTGCTCACTTGGCCACACATCTTCAG 375
Qy 1607 CCAGCCAGGCGCTGCCCAGCTTCAATTACAGACTGACTTCTTCCCTCCCTG 1666
Db 376 CCAGCCAGGCGCTGCCCAGCTTCAATTACAGACTGACTTCTTCCCTCCCTG 435
Qy 1667 CTGTCCAGAGCTGAATGACTTGAATGATGTCACCTGAGTGTCAATGGAGTG 1726
Db 436 CTGTCCAGAGCTGAATGACTTGAATGATGTCACCTGAGTGTCAATGGAGTG 495
Qy 1727 TTATGG 1732
Db 496 TTATGG 501

RESULT 12
US-10-040-862-2721
; Sequence 2721, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
```



```
Db      196 TGGTGAATCTTCCTGGGCTCTGCGCTTGGCCCTGCTCCAGACACATCTTGGCC 255
Qy      1487 TAAGGCTTCTCTCTCAGACCTCTAATTGACCAACAACCACTGGGCTTTCAGCCACAT 1546
Db      256 TAAGGCTTCTCTCTCAGACCTCTAATTGACCAACAACCACTGGGCTTTCAGCCACAT 315
Qy      1547 CAGTGGGCACTGAGAGTGGGGTGCACATGGGGCTGTCTCACTTGGCCACACATCTTCAG 1606
Db      316 CAGTGGGCACTGAGAGTGGGGTGCACATGGGGCTGTCTCACTTGGCCACACATCTTCAG 375
Qy      1607 CCAGCAGAGGCGCTGCGCCAGCTTCAATTTCACAGCTGACTCTTCCCTGCCCCCTG 1666
Db      376 CCAGCAGAGGCGCTGCGCCAGCTTCAATTTCACAGCTGACTCTTCCCTGCCCCCTG 435
Qy      1667 CTGTCCAGAGCTGAACATAGACTTTCGACTTGATGTCACTGGAGTGCACATGGAGTG 1726
Db      436 CTGTCCAGAGCTGAACATAGACTTTCGACTTGATGTCACTGGAGTGCACATGGAGTG 495
Qy      1727 TTATGG 1732
Db      496 TTATGG 501
```

RESULT 14

```
US-10-154-884B-2721
; Sequence 2721, Application US/10154884B
```

```
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane A.
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2721
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(523)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-2721
```

Query Match

23.1%; Score 435; DB 17; Length 523;

```
Best Local Similarity 99.8%; Pred. No. 1,4e-214;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1247 AGACTCTCTTACCCCCAGAGTGTGCCCCCTTAACTTCGAGAGCGGGGCTGGCCCTGTA 1306
Db      16 AGACTCTCTTACCCCCAGAGTGTGCCCCCTTAACTTCGAGAGCGGGGCTGGCCCTGTA 75
Qy      1307 TCTCAGAAAGGAGGGGCAACAGCTTACACCTTCAACAAAGGCCCTTCACATTTGATCTC 1366
Db      76 TCTCAGAAAGGAGGGGCAACAGCTTACACCTTCAACAAAGGCCCTTCACATTTGATCTC 135
Qy      1367 TGAATCTTGGCTGTCTGCACTGTCAAGGTGACACACTGAGCTCATGCTGACACTGCCCC 1426
Db      136 TGAATCTTGGCTGTCTGCACTGTCAAGGTGACACACTGAGCTCATGCTGACACTGCCCC 195
Qy      1427 TGGTGAATCTTCCCTGGGCTGTGCGCTTGGCCCTGCTTCCAGACACACTTCTTGGCC 1486
Db      196 TGGTGAATCTTCCCTGGGCTGTGCGCTTGGCCCTGCTTCCAGACACACTTCTTGGCC 255
Qy      1487 TAAGGCTTCTCTCTCAGACCTCTAATTGACCAACAACCACTGGGCTTTCAGCCACAT 1546
Db      256 TAAGGCTTCTCTCTCAGACCTCTAATTGACCAACAACCACTGGGCTTTCAGCCACAT 315
Qy      1547 CAGTGGGCACTGAGAGTGGGGTGCACATGGGGCTGTCTCACTTGGCCACACATCTTCAG 1606
Db      316 CAGTGGGCACTGAGAGTGGGGTGCACATGGGGCTGTCTCACTTGGCCACACATCTTCAG 375
Qy      1607 CCAGCAGAGGCGCTGCGCCAGCTTCAATTTCACAGCTGACTCTTCCCTGCCCCCTG 1666
Db      376 CCAGCAGAGGCGCTGCGCCAGCTTCAATTTCACAGCTGACTCTTCCCTGCCCCCTG 435
Qy      1667 CTGTCCAGAGCTGAACATAGACTTTCGACTTGATGTCACTGGAGTGCACATGGAGTG 1726
Db      436 CTGTCCAGAGCTGAACATAGACTTTCGACTTGATGTCACTGGAGTGCACATGGAGTG 495
Qy      1727 TTATGG 1732
Db      496 TTATGG 501
```

RESULT 15

```
US-10-764-324-2721
; Sequence 2721, Application US/10764324
```

```
; Publication No. US2004017539A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/764,324
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
```

```
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2721
/ LENGTH: 523
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (15)
/ OTHER INFORMATION: n=A,T,C or G
/ NAME/KEY: unsure
/ LOCATION: (502)
/ OTHER INFORMATION: n=A,T,C or G
US-10-764-324-2721
```

```
Query Match      23.1%; Score 435; DB 18; Length 523;
Best Local Similarity 99.8%; Pred. No. 1.4e-214;
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1247 AAGCTCTCTTACCCCGAGGTGCCCCCTTACGCTCCGAGGCGGGGCTGGGCTGTG 1306
DB 16 AGACTCTCTTACCCCGAGGTGCCCCCTTACGCTCCGAGGCGGGGCTGGGCTGTG 75
QY 1307 TCTCAAGAGGAGGGGACAGCTACACCTACCAAGGCCCCCTGCACTTATATTC 1366
DB 76 TCTCAAGAGGAGGGGACAGCTACACCTACCAAGGCCCCCTGCACTTATATTC 135
QY 1367 TGAATCTGGGCTGTCTGCACTGTACAGGTGCAACACTGCTCATGTCTCACATGCC 1426
DB 136 TGAATCTGGGCTGTCTGCACTGTACAGGTGCAACACTGCTCATGTCTCACATGCC 195
QY 1427 TGTGTGATCTTCTCCCTGGGCTTGCCTGGGCTGCTCCAGACACTTTTGGGC 1486
DB 196 TGTGTGATCTTCTCCCTGGGCTTGCCTGGGCTGCTCCAGACACTTTTGGGC 255
QY 1487 TAAAGGCTTCTCTCTAGAGACCTCTAATTGACCAAGCAACCTGGGCTTCAAGCAAT 1546
DB 256 TAAAGGCTTCTCTCTAGAGACCTCTAATTGACCAAGCAACCTGGGCTTCAAGCAAT 315
QY 1547 CAGTGGGACCTGAGGTGGGGTGCACATGGGGCTGCTCACTTGGCCACACTTCCAG 1606
DB 316 CAGTGGGACCTGAGGTGGGGTGCACATGGGGCTGCTCACTTGGCCACACTTCCAG 375
QY 1607 CCAAGCCAGGGCCCTGCGCCAGCTTCAATTACAGACTGACTGCTTCACTTCCCTG 1666
DB 376 CCAAGCCAGGGCCCTGCGCCAGCTTCAATTACAGACTGACTGCTTCACTTCCCTG 435
QY 1667 CTGTCCAGAGCTGAATAGACTTGGACTTGGATGTCACTGAGTGTCAATGGAGTG 1726
DB 436 CTGTCCAGAGCTGAATAGACTTGGACTTGGATGTCACTGAGTGTCAATGGAGTG 495
QY 1727 TTATGG 1732
DB 496 TTATGG 501
```

```
RESULT 16
US-10-037-270-853
/ Sequence 853, Application US/10037270
/ Publication No. US20030104529A1
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyun
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
```

```
/ APPLICANT: Xue, Aldong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yunding
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: Tillinghast, John
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/10/037,270
/ PRIOR FILING DATE: 2002-01-04
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 853
/ LENGTH: 1067
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (151) .. (984)
US-10-037-270-853
```

```
Query Match      19.5%; Score 368; DB 15; Length 1067;
Best Local Similarity 99.4%; Pred. No. 7.1e-180;
Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 37 AACAGTGGCTGGGGCGGCGGCGGAGTCAACGACCCCAATCTGAGCGGTTCTG 96
DB 64 AACAGTGGCTGGGGCGGCGGCGGAGTCAACGACCCCAATCTGAGCGGTTCTG 123
QY 97 AGCTTGTTCGCGCTCCCTCCCGGGAATGGCGCTATCCGGGTGACCCCGGCGGTC 156
DB 124 AGCTTGTTCGCGCTCCCTCCCGGGAATGGCGCTATCCGGGTGACCCCGGCGGTC 183
QY 157 TGGAGAGAGATGATGCTCTGCACTATCTACGGGATGCTGTCTGCTTATGTTGAG 216
DB 184 TGGAGAGAGATGATGCTCTGCACTATCTACGGGATGCTGTCTGCTTATGTTGAG 243
QY 217 GTGGTGGGGGGAATGAACCGAGTGCAGCTGAGCTCTGAGCTTGTGCTGATGAG 276
DB 244 GTGGTGGGGGGAATGAACCGAGTGCAGCTGAGCTCTGAGCTTGTGCTGATGAG 303
QY 277 GCTCTGGGCGCGCCGAGGCTTAAAGCCCGGCGCGGAGGCTCTGCTGAG 336
DB 304 GCTCTGGGCGCGCCGAGGCTTAAAGCCCGGCGCGGAGGCTCTGCTGAG 363
QY 337 CTGGAAGCGCGGCGGAGTGGCGGAGCAACCTGCGGCTGTGGGGCACTCTGCGCC 396
DB 364 CTGGAAGCGCGGCGGAGTGGCGGAGCAACCTGCGGCTGTGGGGCACTCTGCGCC 423
QY 397 GTGCTGGCCCGGACAGACTGCTGCGGCACTGAGCGCGCAAGGCGGCGGCAAGTGTCT 456
DB 424 GTGCTGGCCCGGACAGACTGCTGCGGCACTGAGCGCGCAAGGCGGCGGCAAGTGTCT 483
QY 457 CCAGAACGCTATAGCTATGAGCACTTCAAGCTTCAAGAGGACAGAGGATAGCTCCGT 516
DB 484 CCAGAACGCTATAGCTATGAGCACTTCAAGCTTCAAGAGGACAGAGGATAGCTCCGT 543
QY 517 CGCCGTGGGAGTCAAGCAAGTCTGCAATTCTCAGCAGG 557
DB 544 CGCCGTGGGAGTCAAGCAAGTCTGCAAAATCTCAGCAGG 584
```

```
RESULT 17
US-10-117-722-853
/ Sequence 853, Application US/10117722
```


NAME/KEY: misc feature
LOCATION: (1)...(451)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-28421

Query Match 15.9%; Score 300; DB 10; Length 451;
Best Local Similarity 100.0%; Pred. No. 1.4e-144;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 882 CTGAGAGCTCTCTATCTGAGAGGCTTTGGGGGAGCTACTGAGTGGGCGCCCTGCTGCA 941
DB 42 CTGAGAGCTCTCTATCTGAGAGGCTTTGGGGGAGCTACTGAGTGGGCGCCCTGCTGCA 101
QY 942 GGCCCTGGGGGAGGCTGTTCTCTGCTGAGAGGCTTGGGAGAGGCTGTGGGAGGAGGCTGT 1001
DB 102 GGCCCTGGGGGAGGCTGTTCTCTGCTGAGAGGCTTGGGAGAGGCTGTGGGAGGAGGCTGT 161
QY 1002 TCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061
DB 162 TCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 221
QY 1062 GATGAT 1121
DB 222 GATGAT 281
QY 1122 TGATCCCACTCCCAAGTCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1181
DB 282 TGATCCCACTCCCAAGTCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 341

RESULT 20

US-10-001-254-7
Sequence 7, Application US/10001254
Publication No. US20030049702A1

GENERAL INFORMATION:
APPLICANT: Reed, John C.
APPLICANT: Godzik, Adam
APPLICANT: Pawlowski, Krzysztof
APPLICANT: Fiorentino, Loredana
APPLICANT: Lee, Sung Hyung
APPLICANT: Roth, Wilfried
APPLICANT: Stemmer-Bliwien, Frank
TITLE OF INVENTION: No. US20030049702A1e1 Death Domain Proteins
FILE REFERENCE: P-1J 5037
CURRENT APPLICATION NUMBER: US/10/001.254
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/301,889
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/715,893
PRIOR FILING DATE: 2000-11-17
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 303
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(303)
US-10-001-254-7

Query Match 13.4%; Score 252; DB 14; Length 303;
Best Local Similarity 99.7%; Pred. No. 1.1e-119;

Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 157 TGGAGAGAGATGATGCTCTGAGCTACTAGAGAGTGTGTGCTTCAACCGTATGTCAG 216
DB 1 TGGAGAGAGATGATGCTCTGAGCTACTAGAGAGTGTGTGCTTCAACCGTATGTCAG 60
QY 217 GTGTGGGGGGGCACTGAT 276
DB 61 GTGTGGGGGGGCACTGAT 120

QY 277 GCTCCTGAGCGCGCCGAGAGCTTAGCCCGGCGCCGAGCGGCTAGAGCTCTGCTGAG 336
DB 121 GCTCCTGAGCGCGCCGAGAGCTTAGCCCGGCGCCGAGCGGCTAGAGCTCTGCTGAG 180
QY 337 CTGAGAGCGCGGAT 396
DB 181 CTGAGAGCGCGGAT 240
QY 397 GTGCTGGCGCGCAGAT 456
DB 241 GTGCTGGCGCGCAGAT 300
QY 457 CCA 459
DB 301 CCA 303

RESULT 21

US-09-867-701-3069
Sequence 3069, Application US/09867701
Patent No. US20020132237A1

GENERAL INFORMATION:
APPLICANT: Agiate, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.497
CURRENT APPLICATION NUMBER: US/09/867,701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3069
LENGTH: 441
TYPE: DNA
ORGANISM: Homo sapien
US-09-867-701-3069

Query Match 12.6%; Score 237; DB 9; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.3e-112;

Matches 237; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1209 TCTGAGGCTTAAACATTTCCAGCTGAGTTCTTCCAGACTCTCTAACCAGAGTG 1268
DB 199 TCTGAGGCTTAAACATTTCCAGCTGAGTTCTTCCAGACTCTCTAACCAGAGTG 258
QY 1269 TGCCTCTTGAAGCTCCGAGAGGCGGGGCTGGGCTGTATCTCAGAGGAGGCGACAGC 1328
DB 259 TGCCTCTTGAAGCTCCGAGAGGCGGGGCTGGGCTGTATCTCAGAGGAGGCGACAGC 318
QY 1329 TCACAGCTACCAAGAGGCGGCGCTGACATTTGATCTGATCTTGGGCTGTGCACTG 1388
DB 319 TCACAGCTACCAAGAGGCGGCGCTGACATTTGATCTTGGGCTGTGCACTG 378
QY 1389 TCACAGCTACCAAGAGGCGGCGCTGACATTTGATCTTGGGCTGTGCACTG 1445
DB 379 TCACAGCTACCAAGAGGCGGCGCTGACATTTGATCTTGGGCTGTGCACTG 435

RESULT 22

US-09-918-995-12268
Sequence 12268, Application US/09918995
Publication No. US20030073623A1

GENERAL INFORMATION:
APPLICANT: Hyeeq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12268
; LENGTH: 484
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(484)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-12268

Query Match
Best Local Similarity 100.0%; Score 96; DB 10; Length 484;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 873 CATCAAGTCTTCAGAGCTTCCTATCTGAGAGCCTTTGGGGGAGCACTAGTGGCGC 932
DB 36 CATCAAGTCTTCAGAGCTTCCTATCTGAGAGCCTTTGGGGGAGCACTAGTGGCGC 95

QY 933 CTTGCTGAGAGCCCTGGGGGCGGTCTCTGACTGA 968
DB 96 CTTGCTGAGAGCCCTGGGGGCGGTCTCTGACTGA 131

RESULT 23
US-10-425-115-11108
; Sequence 11108, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 11108
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_32821C.1
US-10-425-115-11108

Query Match
Best Local Similarity 100.0%; Score 55; DB 18; Length 180;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1127 CCACCTTCAGAGTCTCCGGGCAACCTTCTCTGGAGAGACACCATCTTACCCCT 1181
DB 84 CCACCTTCAGAGTCTCCGGGCAACCTTCTCTGGAGAGACACCATCTTACCCCT 138

RESULT 24
US-10-037-270-853/C
; Sequence 853, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
```

```

; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dramanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1e1 Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 853
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(984)
US-10-117-722-853

Query Match
Best Local Similarity 100.0%; Score 37; DB 15; Length 1067;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1760 ACATGGGGCCAAACAGTAAACAGCCACCTTCTTGG 1796
DB 979 ACATGGGGCCAAACAGTAAACAGCCACCTTCTTGG 943

RESULT 25
US-10-117-722-853/C
; Sequence 853, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Dramanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1e1 Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 853
; LENGTH: 1067
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(984)
US-10-117-722-853

Query Match
Best Local Similarity 100.0%; Score 37; DB 17; Length 1067;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1760 ACATGGGGCCAAACAGTAAACAGCCACCTTCTTGG 1796
DB 979 ACATGGGGCCAAACAGTAAACAGCCACCTTCTTGG 943
```

```
RESULT 26
US-09-814-353-18768/C
; Sequence 18768, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18768
; LENGTH: 838
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 100, 101, 125, 126, 128, 129, 131, 138, 164, 174, 175, 176,
; LOCATION: 184, 185
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-18768

Query Match      1.2%; Score 22; DB 10; Length 838;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      158 GGGAGGAGTGTGTCCTGGA 179
Db      381 GGGAGGAGTGTGTCCTGGA 360

RESULT 27
US-10-343-710-27
; Sequence 27, Application US/10343710
; Publication No. US20040087478A1
; GENERAL INFORMATION:
; APPLICANT: GILLEN, Clemens
; APPLICANT: WETZELS, Ingrid
; APPLICANT: WENDT, Stephan
; APPLICANT: WEIHE, E.
; APPLICANT: SCHAEFER, M., K.-H.
; TITLE OF INVENTION: SCREENING METHOD
; FILE REFERENCE: 029310.52022US
; CURRENT APPLICATION NUMBER: US/10/343,710
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/EP01/09011
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 157
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 27
; LENGTH: 1339
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
```

```
; NAME/KEY: misc_feature
; LOCATION: (338)..(338)
; OTHER INFORMATION: n is a, c, g, or t
US-10-343-710-27

Query Match      1.2%; Score 22; DB 17; Length 1339;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1449 CTGCCTGGCTGCTTCCGAGC 1470
Db      1018 CTGCCTGGCTGCTTCCGAGC 1039

RESULT 28
US-10-001-254-47
; Sequence 47, Application US/10001254
; Publication No. US20030049702A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; APPLICANT: Pawlowski, Frank
; APPLICANT: Fiorentino, Loredana
; APPLICANT: Lee, Sug Hyung
; APPLICANT: Roth, Wilfred
; APPLICANT: Stemmer-Liwen, Frank
; TITLE OF INVENTION: P-Lf 5037
; FILE REFERENCE: P-Lf 5037
; CURRENT APPLICATION NUMBER: US/10/001,254
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/301,889
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/715,893
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic primer
US-10-001-254-47

Query Match      1.1%; Score 21; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      56 CGCCGAGATCATCGAGCGC 76
Db      1 CGCCGAGATCATCGAGCGC 21

RESULT 29
US-09-814-353-13540
; Sequence 13540, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
```

```

; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13540
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-814-353-13540

Query Match          1.1%; Score 21; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      705 GACCTGTGACATCCGGCTCCG 725
DB      279 GACCTGTGACATCCGGCTCCG 299

RESULT 30
US-10-425-114-827/c
; Sequence 827, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Yindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 827
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700150477_FLI
US-10-425-114-827

Query Match          1.1%; Score 21; DB 17; Length 1439;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      594 GCGGCGGAGTCGGGGCCGCC 614
DB      171 GCGGCGGAGTCGGGGCCGCC 151

RESULT 31
US-10-425-115-14144/c
; Sequence 1414, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 14144
```

```

; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_11288C.1
US-10-425-115-14144

Query Match          1.1%; Score 21; DB 18; Length 1496;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      594 GCGGCGGAGTCGGGGCCGCC 614
DB      173 GCGGCGGAGTCGGGGCCGCC 153

RESULT 32
US-10-437-963-9016/c
; Sequence 9016, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 9016
; LENGTH: 2296
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15472C.1
US-10-437-963-9016

Query Match          1.1%; Score 21; DB 18; Length 2296;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1042 GCGGCGGCGCCTGTGCTG 1062
DB      248 GCGGCGGCGCCTGTGCTG 228

RESULT 33
US-10-305-720-1432/c
; Sequence 1432, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expressi
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1432
; LENGTH: 3234
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank ID No. US20040010136A1 g599819
```

US-10-305-720-1432

Query Match 1.1%; Score 21; DB 17; Length 3234;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1058 TGCTGATGAGAGAGAGGGG 1078

Db 2746 TGCTGATGAGAGAGAGAGGGG 2726

RESULT 34

US-10-087-192-1045/c

; Sequence 1045, Application US/10087192

; Publication No. US20020182586A1

; GENERAL INFORMATION:

; APPLICANT: Morrie, David W.

; APPLICANT: Engelhard, Eric K.

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR

; FILE REFERENCE: 529452000122

; CURRENT APPLICATION NUMBER: US/10/087,192

; CURRENT FILING DATE: 2002-03-01

; PRIOR APPLICATION NUMBER: US 09/747,377

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 09/798,586

; PRIOR FILING DATE: 2001-03-02

; NUMBER OF SEQ ID NOS: 2059

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1045

; LENGTH: 24902

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(24902)

; OTHER INFORMATION: n = A,T,C or G

US-10-087-192-1045

Query Match 1.1%; Score 21; DB 13; Length 24902;

Best Local Similarity 100.0%; Pred. No. 4.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1660 CCCCTGCTGTCGAGAGCTGA 1680

Db 19314 CCCCTGCTGTCGAGAGCTGA 19294

RESULT 35

US-10-425-115-46300/c

; Sequence 46300, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 46300

; LENGTH: 411

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_142231C.1

US-10-425-115-46300

Query Match 1.1%; Score 20; DB 18; Length 411;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 411 CGACCTGCTGCGGCACTGGA 430

Db 154 CGACCTGCTGCGGCACTGGA 135

RESULT 36

US-10-437-963-90512/c

; Sequence 90512, Application US/10437963

; Publication No. US2004012343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 90512

; LENGTH: 424

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_89175C.1

US-10-437-963-90512

Query Match 1.1%; Score 20; DB 18; Length 424;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 629 GACGCGCGCGAGAGGCGCC 648

Db 101 GACGCGCGCGAGAGGCGCC 82

RESULT 37

US-10-029-386-11816/c

; Sequence 11816, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 11816

; LENGTH: 540

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO CHR18.3

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7

; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3

; OTHER INFORMATION: NT HIT: L33260.1, EVALUE 8.00e+00

; OTHER INFORMATION: SWISSPROT HIT: P38335, EVALUE 3.90e+00

; OTHER INFORMATION: EST_HUMAN HIT: AA62107.1, EVALUE 7.60e-01

US-10-029-386-11816

Query Match 1.1%; Score 20; DB 16; Length 540;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1232 CTGAGTTTCTCTCCAGACT 1251
DB 493 CTGAGTTTCTCTCCAGACT 474

RESULT 38

US-10-282-596-42/c
; Sequence 42, Application US/10282596
; Publication No. US2003023632A1
; GENERAL INFORMATION:
; APPLICANT: Erlander, Mark
; APPLICANT: Ma, Xiao-Jun
; APPLICANT: Sgroi, Dennis C.
; TITLE OF INVENTION: BIOMARKERS FOR BREAST CANCER
; FILE REFERENCE: 47577-20043.21
; CURRENT APPLICATION NUMBER: US/10/282,596
; CURRENT FILING DATE: 2002-12-28
; PRIOR APPLICATION NUMBER: US 10/028,018
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 10/211,015
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1066
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Clone ID 5768283
US-10-282-596-42

Query Match 1.1%; Score 20; DB 17; Length 1066;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1340 CAAAGGCCCCCTGCACATT 1359
DB 1051 CAAAGGCCCCCTGCACATT 1032

RESULT 39

US-10-156-761-6280/c
; Sequence 6280, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, TUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6280
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2304)
US-10-156-761-6280

Query Match 1.1%; Score 20; DB 15; Length 2304;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 344 GCCGCGGCGAGTCGCGGAG 363
DB 2258 GCCGCGGCGAGTCGCGGAG 2239

RESULT 40

US-10-246-330-1
; Sequence 1, Application US/10246330
; Publication No. US20030166030A1
; GENERAL INFORMATION:
; APPLICANT: O'Toole, George A.
; APPLICANT: Mah, Thien-Pah
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF
; FILE REFERENCE: 14537-002001
; CURRENT APPLICATION NUMBER: US/10/246,330
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,241
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2610
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2607)
US-10-246-330-1

Query Match 1.1%; Score 20; DB 16; Length 2610;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 CGCGCAAGCGCGCGGCCA 450
DB 1208 CGCGCAAGCGCGCGGCCA 1227

RESULT 41

US-10-087-192-1279
; Sequence 1279, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David M.
; APPLICANT: Engelhardt, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1279
; LENGTH: 31812
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(31812)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1279

Query Match 1.1%; Score 20; DB 13; Length 31812;

Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1799 AGGGATGCAAGGCTTTGG 1818
DB 18865 AGGGATGCAAGGCTTTGG 18884

RESULT 42

US-10-723-860-2504/C
; Sequence 2504, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2504
; LENGTH: 115863
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-2504

Query Match 1.1%; Score 20; DB 18; Length 115863;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1244 CCCGACTCTCTACCCCC 1263
DB 73742 CCCGACTCTCTACCCCC 73723

RESULT 43

US-10-292-798-1369/C
; Sequence 1369, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1369
; LENGTH: 744802
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(744802)
; NAME/KEY: CDS
; LOCATION: (201)..(246)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25640)..(25677)

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27078)..(27094)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (141192)..(141769)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (159571)..(159606)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (174525)..(174575)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (234891)..(235013)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (235514)..(235560)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (279677)..(279729)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (408660)..(409123)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409204)..(409669)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (428381)..(428396)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (472204)..(472330)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (714252)..(714355)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (714447)..(714529)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (739794)..(739881)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (744484)..(744602)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (51812)..(51911)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (57122)..(57221)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (79368)..(79467)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (310089)..(310188)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (332935)..(332935)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (332992)..(332992)
; OTHER INFORMATION: a, t, c, g, unknown or other


```

FEATURE:
NAME/KEY: modified base
LOCATION: (362002)..(362101)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (639781)..(639880)
OTHER INFORMATION: a, t, c, g, unknown or other
US-10-292-798-1369

```

```

Query Match      1.1%; Score 20; DB 17; Length 744802;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1199 GGATGGGCTCTGAGCCCT 1218
      |||||
DB      238847 GGATGGGCTCTGAGCCCT 238828

```

```

RESULT 44
US-10-156-761-1

```

```

Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: misc feature
LOCATION: (4187715)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

```

```

Query Match      1.1%; Score 20; DB 15; Length 9025608;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      344 GCGCGGCGAGTCGCGGCGAG 363
      |||||
DB      7567115 GCGCGGCGAGTCGCGGCGAG 7567134

```

```

RESULT 45
US-09-908-975-11385
Sequence 11385, Application US/09908975
Publication No. US20030165843A1
GENERAL INFORMATION:
APPLICANT: SHOSHANI, AVI
APPLICANT: WASSERMAN, ALON
APPLICANT: MINTZ, ELLI
APPLICANT: MINTZ, LIAT
APPLICANT: FAIGER, SIMCHON
TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING
FILE REFERENCE: 36688-0005
CURRENT APPLICATION NUMBER: US/09/908,975

```

```

CURRENT FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: US 60/287,724
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: US 60/221,607
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 32337
SOFTWARE: PatentIn version 3.0
SEQ ID NO 11385
LENGTH: 60
TYPE: DNA
ORGANISM: Homo sapiens
US-09-908-975-11385

```

```

Query Match      1.0%; Score 19; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1017 TGTGATGAGGCTGACTAT 1035
      |||||
DB      6 TGTGATGAGGCTGACTAT 24

```

```

RESULT 46
US-10-719-993-13899/c

```

```

Sequence 13899, Application US/10719993
Publication No. US20040265849A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001496
CURRENT APPLICATION NUMBER: US/10/719,993
CURRENT FILING DATE: 2003-11-24
NUMBER OF SEQ ID NOS: 55342
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13899
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-719-993-13899

```

```

Query Match      1.0%; Score 19; DB 18; Length 201;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      485 GCTCTTCAAGAGGACAGA 503
      |||||
DB      33 GCTCTTCAAGAGGACAGA 15

```

```

RESULT 47
US-10-432-422-189/c
Sequence 189, Application US/10432422
Publication No. US20040076981A1
GENERAL INFORMATION:
APPLICANT: Syngenta Participations AG
APPLICANT: Cornell Research Foundation, Inc.
APPLICANT: Yoder, Olen
APPLICANT: Turgeon, Barbara G.
APPLICANT: Lu, Shen-wen
TITLE OF INVENTION: Fungal Iron Reductase Gene
FILE REFERENCE: 1360.017M01
CURRENT APPLICATION NUMBER: US/10/432,422
CURRENT FILING DATE: 2003-05-21
PRIOR APPLICATION NUMBER: US 60/252,732
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/252,649
PRIOR FILING DATE: 2000-11-22
NUMBER OF SEQ ID NOS: 210
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 189
LENGTH: 327
TYPE: DNA

```

ORGANISM: Cochliobolus heterostrophus
US-10-432-422-189

Query Match 1.0%; Score 19; DB 17; Length 327;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1062 GATGAGAGAGAGAGGGG 1080
DB 157 GATGAGAGAGAGAGGGGG 139

RESULT 48
US-09-983-965-824
Sequence 824, Application US/09983965
Patent No. US20020137160A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
PRIOR FILING DATE: 2001-10-26
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 824
LENGTH: 397
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
OTHER INFORMATION: Clone ID: 28-LIB188-014-Q1-E1-G7
US-09-983-965-824

Query Match 1.0%; Score 19; DB 9; Length 397;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TTGTCCGCTCCCTCCCC 118
DB 60 TTGTCCGCTCCCTCCCC 78

RESULT 49
US-10-425-115-88370
Sequence 88370, Application US/10425115
Publication No. US2004024272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
PRIOR FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 88370
LENGTH: 400
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_180594C.1
US-10-425-115-88370

Query Match 1.0%; Score 19; DB 18; Length 400;
Best Local Similarity 100.0%; Pred. No. 67;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1079 GGCGGGCCCCGACAGAGC 1097
DB 186 GGCGGGCCCCGACAGAGC 204

RESULT 50
US-09-864-761-1941
Sequence 1941, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmlca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 1941
LENGTH: 452
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035427.17
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
US-09-864-761-1941

Query Match 1.0%; Score 19; DB 9; Length 452;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1401 ACACTGGCTCATGCTCACA 1419
Db 113 ACACTGGCTCATGCTCACA 131

Search completed: March 24, 2005, 07:34:41
Job time : 1711.81 secs

THIS PAGE LEFT BLANK

ALIGNMENTS

RESULT 1
US-09-023-655-182

/ Sequence 182, Application US/09023655
/ Patent No. 6607879
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Selhammer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HERewith
/ CLASSIFICATION:
/ PRIORITY APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0001 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 182:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 847 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: FIBROTO1
/ CLONE: 053532
/ US-09-023-655-182

Query Match 37.5%; Score 707; DB 4; Length 847;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1177 CCCCTTGACAGCCCTCCACAGATGTGGGCTCTGAGGCTTAACCATTTCCAGCTGAG 1236
DB 141 CCCCTTGACAGCCCTCCACAGATGTGGGCTCTGAGGCTTAACCATTTCCAGCTGAG 200
QY 1237 TTTCCTTCCAGACTCCTCTACCCCGAGGTGTGCCCTTTAGCCTTCGGAGGCGGGGGC 1296
DB 201 TTTCCTTCCAGACTCCTCTACCCCGAGGTGTGCCCTTTAGCCTTCGGAGGCGGGGGC 260
QY 1297 TGGGCTGTATCTCAAGAGGAGGGGCAAGCTACACACTCAACAAAGGCCCTTCGAC 1356
DB 261 TGGGCTGTATCTCAAGAGGAGGGGCAAGCTACACACTCAACAAAGGCCCTTCGAC 320
QY 1357 ATTGTATCTGTATCTTGGGCTGTGCACTGTCAAGAGGCAACACTGCTCATGCTC 1416
DB 321 ATTGTATCTGTATCTTGGGCTGTGCACTGTCAAGAGGCAACACTGCTCATGCTC 380

QY 1417 ACACTGCCCTGCTGATGATCTTCCCTGCGGCTCTGCGCTTGTCCAGACACAC 1476
DB 381 ACACTGCCCTGCTGATGATCTTCCCTGCGGCTCTGCGCTTGTCCAGACACAC 440
QY 1477 TTCTTTGGCTTAAGGGCTTCTCTGAGGACCTCTAATTGACCAACCAACTGGGCT 1536
DB 441 TTCTTTGGCTTAAGGGCTTCTCTGAGGACCTCTAATTGACCAACCAACTGGGCT 500
QY 1537 TCAGCCACATAGTGGGCACTGAGCTGGGGTGACATGGGGCTGCTACCTTGGCCAC 1596
DB 501 TCAGCCACATAGTGGGCACTGAGCTGGGGTGACATGGGGCTGCTACCTTGGCCAC 560
QY 1597 ACATCTTCAGCCAGCCAGGGCCCTGCCAGCTTCAATTACAGACTGACTCTTCACC 1656
DB 561 ACATCTTCAGCCAGCCAGGGCCCTGCCAGCTTCAATTACAGACTGACTCTTCACC 620
QY 1657 TTCCCCCTGCTGTCCAGAGCTGAACATGACTTGACTTGATGTCACCTGGAGTGA 1716
DB 621 TTCCCCCTGCTGTCCAGAGCTGAACATGACTTGACTTGATGTCACCTGGAGTGA 680
QY 1717 CATGGAGTGTATGACAGATCATACCAAGGCTTACTGTGACATGGGGCCCAACCA 1776
DB 681 CATGGAGTGTATGACAGATCATACCAAGGCTTACTGTGACATGGGGCCCAACCA 740
QY 1777 GTAAACAGCCACTTCTTGGAAAAGGAATGCAAAAGCTTTGGGGGTGATGAAAAGACT 1836
DB 741 GTAAACAGCCACTTCTTGGAAAAGGAATGCAAAAGCTTTGGGGGTGATGAAAAGACT 800
QY 1837 TTACCAATGATACCAATTAACCTGCGCTGGAAGGGCATAGGTGG 1883
DB 801 TTACCAATGATACCAATTAACCTGCGCTGGAAGGGCATAGGTGG 847

RESULT 2

US-09-620-312D-853

/ Sequence 853, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jjian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yunging
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_ft_genes Version 1.0
/ SEQ ID NO 853
/ LENGTH: 1067
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS

LOCATION: (151)..(984)
US-09-620-312D-853

Query Match 19.5%; Score 368; DB 4; Length 1067;
Best Local Similarity 99.4%; Pred. No. 1.6e-167;
Matches 518; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 37 AACAGTGGCTGCGCCGCGCCAGAGTATCGAGCGCAGAACTGGCCGGGTTCTG 96
DB 64 AACAGTGGCTGCGCCGCGCGCCAGAGTATCGAGCGCAGAACTGGCCGGGTTCTG 123
QY 97 AGCTTTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 156
DB 124 AGCTTTGTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183
QY 157 TGGGAGAGGATGATGCTGCTGATCTAGCGGAGTGTGTGCTTACCCGATGTTGAG 216
DB 184 TGGGAGAGGATGATGCTGCTGATCTAGCGGAGTGTGTGCTTACCCGATGTTGAG 243
QY 217 GTGGTGGCGCGGCACTGACCGAGTGGAGCTGCTGCGCTTCTGTGATGATGAG 276
DB 244 GTGGTGGCGCGGCACTGACCGAGTGGAGCTGCTGCGCTTCTGTGATGATGAG 303
QY 277 GCTCTGCGCGCGCGCGGAGGCTTACCGCGCGCGCGCGCTTACCGCTTCTGAG 336
DB 304 GCTCTGCGCGCGCGCGGAGGCTTACCGCGCGCGCGCGCTTACCGCTTCTGAG 363
QY 337 CTGAGAGCGCGCGCGCGGAGTGGCGGAGAGCACTGTGCGCTTCTGTGAGGCTG 396
DB 364 CTGAGAGCGCGCGCGCGGAGTGGCGGAGAGCACTGTGCGCTTCTGTGAGGCTG 423
QY 397 GTCTGGCGCGCGCGCGCACTGCTGCGCACTGTGCGCGCGCGCGCGCGCGCTG 456
DB 424 GTCTGGCGCGCGCGCGCACTGCTGCGCACTGTGCGCGCGCGCGCGCGCGCTG 483
QY 457 CCAGAACGCTATAGCTATGAGCACTTCCAGCTTTCAAGAGAGAGAGGTAAGTGC 516
DB 484 CCAGAACGCTATAGCTATGAGCACTTCCAGCTTTCAAGAGAGAGAGGTAAGTGC 543
QY 517 CGCGCTGGGAGTCAAGCACTTTCGCAAAATTCACAGAGG 557
DB 544 CGCGCTGGGAGTCAAGCACTTTCGCAAAATTCACAGAGG 584
```

RESULT 3

US-09-620-312D-853/c
Sequence 853, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Lau, Chenghua
APPLICANT: Auandi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungling
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Dimanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: 784CTP28
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_Fl_genes Version 1.0
SEQ ID NO 853
LENGTH: 1067
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (151)..(984)
US-09-620-312D-853

Query Match 2.0%; Score 37; DB 4; Length 1067;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1760 ACATGGGGCCAAACAGTAAACAGCCACTTCTTGG 1796
DB 979 ACATGGGGCCAAACAGTAAACAGCCACTTCTTGG 943
```

RESULT 4

US-09-016-434-1432/c
Sequence 1432, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1432:
SEQUENCE CHARACTERISTICS:
LENGTH: 3234 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9599819
US-09-016-434-1432

Query Match 1.1%; Score 21; DB 4; Length 3234;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1058	TCCTGATGAGGAGCAAGGGG	1078
Db	2746	TGCTGATGAGGAGGAAGGGG	2726

```

RESULT 5
US-09-513-999C-32250
: Sequence 32250, Application US/09513999C
: Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
Patent No. 6783961
FILE REFERENCE: 59. US2. REG
CURRENT APPLICATION NUMBER: US/09/513,999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 32250
LENGTH: 267
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 183
OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-32250

```

```

QY      1826 GGAAGACCTTTACAAT 1845
      |||||
Db      34  GGAAGACCTTTACAAT 53

RESULT 6
US-09-252-991A-2993/c
; Sequence 2993, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 10/7196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2993
; LENGTH: 2673
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2993

```

Query Match	1.1%	Score 20;	DB 4;	Length 2673;
Best Local Similarity	100.0%	Pred. No. 22;		
Matches	20;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;
QY	431	CGCGCAAGCGGCGCGGCCA	450	
Db	1472	CGCGCAAGCGGCGCGGCCA	1453	

```

RESULT 7
US-09-252-991A-2686
; Sequence 2686, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2686
; LENGTH: 2715
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2686

```

```

OY      431  CGCGCAAGCGGCGCCGCCA 450
        |||||||
Db      1313  CGCGCAAGCGGCGCCGCCA 1332

RESULT 8
US-09-252-991A-2875
; Sequence 2875, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2875
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2875

```

```

QY      431  CGCGCAGCGGCGCGGCCA 450
        |||||
Db      1371 CGCGCAGCGGCGCGGCCA 1390

RESULT 9
US-08-787-739-21/c
; Sequence 21, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jarmilr
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
;

```


ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: promoter before transcription start
US-08-787-739-21

Query Match 1.0%; Score 19; DB 3; Length 506;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCTGGCCTGCTCCGAC 1470
DB 215 CCTGGCCTGCTCCGAC 197

RESULT 10
US-08-481-658B-27/c
Sequence 27, Application US/08481658B
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:

ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: Proposed MN promoter
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-481-658B-27

Query Match 1.0%; Score 19; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCTGGCCTGCTCCGAC 1470
DB 215 CCTGGCCTGCTCCGAC 197

RESULT 11
US-08-477-504A-27/c
Sequence 27, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: Proposed MN promoter
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-27

Query Match 1.0%; Score 19; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
|||||
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 12
US-08-486-756A-27/c
Sequence 27, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

DESCRIPTION: Proposed MN promoter
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-27

Query Match 1.0%; Score 19; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
|||||
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 13
US-08-485-862B-27/c
Sequence 27, Application US/08485862B
Patent No. 5989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: Proposed MN promoter
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-27

Query Match 1.0%; Score 19; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
|||||
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 14
US-08-787-739-27/c
; Sequence 27, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; DESCRIPTION: Proposed MN promoter
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-787-739-27

Query Match 1.0%; Score 19; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTCCGAC 1470
Db 215 CCCTGGCCTGCTCCGAC 197

RESULT 15
US-08-487-077A-27/c
; Sequence 27, Application US/08487077A
; Patent No. 6069242
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,077A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 540 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; DESCRIPTION: Proposed MN promoter
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-487-077A-27

Query Match 1.0%; Score 19; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTCCGAC 1470
Db 215 CCCTGGCCTGCTCCGAC 197

RESULT 16
US-08-485-863A-27/c
; Sequence 27, Application US/08485863A
; Patent No. 6093548
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California

COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,863A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3G
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: Proposed MN promoter
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-863A-27

Query Match 1.0%; Score 19; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
|||||
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 17
US-08-485-049D-27/c
Sequence 27, Application US/08485049D
Patent No. 6204370
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,049D
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.

REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: Proposed MN promoter
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-049D-27

Query Match 1.0%; Score 19; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
|||||
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 18
US-09-178-115-27/c
Sequence 27, Application US/09178115
Patent No. 6297041
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
FILE REFERENCE: D-0021.5A
CURRENT APPLICATION NUMBER: US/09/178,115
CURRENT FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 09/177,776
EARLIER FILING DATE: 1998-10-23
EARLIER APPLICATION NUMBER: 08/787,739
EARLIER FILING DATE: 1997-01-24
EARLIER APPLICATION NUMBER: 08/485,049
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/486,756
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/477,504
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/481,658
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,862
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/485,863
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/487,077
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/260,190
EARLIER FILING DATE: 1994-06-15
EARLIER APPLICATION NUMBER: 08/177,093
EARLIER FILING DATE: 1993-12-30
EARLIER APPLICATION NUMBER: 07/964,589
EARLIER FILING DATE: 1992-10-21
EARLIER APPLICATION NUMBER: PV-709-92
EARLIER FILING DATE: 1992-03-11
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 27
LENGTH: 540
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(540)

US-09-178-115-27

Query Match 1.0%; Score 19; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 19

US-09-177-776-27/c
; Sequence 27, Application US/0917776A
; Patent No. 6297051
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; FILE REFERENCE: D-0021.5A
; CURRENT APPLICATION NUMBER: US/09/177,776A
; CURRENT FILING DATE: 1998-10-23
; EARLIER APPLICATION NUMBER: 08/787,739
; EARLIER FILING DATE: 1997-01-24
; EARLIER APPLICATION NUMBER: 08/485,049
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/486,756
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/477,504
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/481,658
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,862
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/485,863
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/487,077
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 08/260,190
; EARLIER FILING DATE: 1994-06-15
; EARLIER APPLICATION NUMBER: 08/177,093
; EARLIER FILING DATE: 1993-12-30
; EARLIER APPLICATION NUMBER: 07/964,589
; EARLIER FILING DATE: 1992-10-21
; EARLIER APPLICATION NUMBER: PV-709-92
; EARLIER FILING DATE: 1992-03-11
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 540
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(540)
US-09-177-776-27

Query Match 1.0%; Score 19; DB 3; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 20
US-09-772-719B-27/c
; Sequence 27, Application US/09772719B
; Patent No. 6770438
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan

Pastorekova, Silvia
Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 465 California Street, Suite 450
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/772,719B
FILING DATE: 30-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3A-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 540 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
DESCRIPTION: Proposed MN promoter
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-772-719B-27

Query Match 1.0%; Score 19; DB 4; Length 540;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
DB 215 CCCTGGCCTGCTTCCGAGC 197

RESULT 21
US-09-949-016-21931/c
; Sequence 21931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21931
; LENGTH: 601

TYPE: DNA
ORGANISM: Human
US-09-949-016-21931

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GCTCTTCAAGAGACAGA 503
DB 584 GCTCTTCAAGAGACAGA 566

RESULT 22
US-09-949-016-21932/c
Sequence 21932, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14, 755
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21932
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-21932

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GCTCTTCAAGAGACAGA 503
DB 233 GCTCTTCAAGAGACAGA 215

RESULT 23
US-09-949-016-21933/c
Sequence 21933, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14, 755
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21933
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-21933

Query Match 1.0%; Score 19; DB 4; Length 601;

Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GCTCTTCAAGAGACAGA 503
DB 93 GCTCTTCAAGAGACAGA 75

RESULT 24
US-09-949-016-21934/c
Sequence 21934, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14, 755
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21934
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-21934

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GCTCTTCAAGAGACAGA 503
DB 77 GCTCTTCAAGAGACAGA 59

RESULT 25
US-09-949-016-21935/c
Sequence 21935, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14, 755
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21935
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-21935

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GCTCTTCAAGAGACAGA 503

Db 30 GCTCTCAAGAGACAGA 12

RESULT 26

US-09-949-016-70295

; Sequence 70295, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70295
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-70295

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AGAGAGGTGCGAGAAGA 31

Db 573 AGAGAGGTGCGAGAAGA 591

RESULT 27

US-09-949-016-79572

; Sequence 79572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79572
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-79572

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1225 TTTCAGCTGAGTTTCCTT 1243

Db 378 TTTCAGCTGAGTTTCCTT 396

RESULT 28

US-09-949-016-79573

; Sequence 79573, Application US/09949016

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79573
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-79573

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1225 TTTCAGCTGAGTTTCCTT 1243

Db 56 TTTCAGCTGAGTTTCCTT 74

RESULT 29

US-09-949-016-79574

; Sequence 79574, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 79574
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-79574

Query Match

Best Local Similarity 100.0%; Score 19; DB 4; Length 601;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1225 TTTCAGCTGAGTTTCCTT 1243

Db 55 TTTCAGCTGAGTTTCCTT 73

RESULT 30

US-09-949-016-84252/c

; Sequence 84252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 84256
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-84256

Query Match 1.0%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 GCTCTTCAAGAGCAGCA 503
DB 30 GCTCTTCAAGAGCAGCA 12

RESULT 35
US-09-270-767-27566
Sequence 27566, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:

APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 27566
LENGTH: 878
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-27566

Query Match 1.0%; Score 19; DB 4; Length 878;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 GCTGGAGCTCTGCGCTT 264
DB 388 GCTGGAGCTCTGCGCTT 406

RESULT 36

US-08-481-658B-58/C
Sequence 58, Application US/08481658B
Patent No. 5955075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (PPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727

INFORMATION FOR SEQ ID NO: 58:

SEQUENCE CHARACTERISTICS:

LENGTH: 904 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-481-658B-58

Query Match 1.0%; Score 19; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCGCTCTCCAGC 1470
DB 642 CCCTGGCGCTCTCCAGC 624

RESULT 37

US-08-477-504A-58/C
Sequence 58, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (PPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 904 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-58

Query Match 1.0%; Score 19; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCTGGCGCTGCTTCCGAGC 1470
|||||
DB 642 CCTGGCGCTGCTTCCGAGC 624

RESULT 38
US-08-486-756A-58/c
; Sequence 58, Application US/08486756A
; Patent No. 5981711
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorek, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,756A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-486-756A-58

Query Match 1.0%; Score 19; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCTGGCGCTGCTTCCGAGC 1470
|||||
DB 642 CCTGGCGCTGCTTCCGAGC 624

RESULT 39
US-08-485-862B-58/c
; Sequence 58, Application US/08485862B
; Patent No. 5989838
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorek, Silvia

APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-485-862B-58

Query Match 1.0%; Score 19; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCTGGCGCTGCTTCCGAGC 1470
|||||
DB 642 CCTGGCGCTGCTTCCGAGC 624

RESULT 40
US-08-487-077A-58/c
; Sequence 58, Application US/08487077A
; Patent No. 6069242
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorek, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,077A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3H
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 904 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-077A-58

Query Match
Best Local Similarity 100.0%; Score 19; DB 3; Length 904;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
DB 642 CCCTGGCCTGCTTCCGAGC 624

RESULT 41
US-08-485-863A-58/c
Sequence 58, Application US/08485863A
Patent No. 6093548
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSER: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,863A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3G
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 904 base pairs

```

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-863A-58

Query Match
Best Local Similarity 100.0%; Score 19; DB 3; Length 904;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
DB 642 CCCTGGCCTGCTTCCGAGC 624

RESULT 42
US-08-485-049D-58/c
Sequence 58, Application US/08485049D
Patent No. 6204370
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSER: Leona L. Lauder
STREET: 369 Pine Street
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,049D
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 904 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-049D-58

Query Match
Best Local Similarity 100.0%; Score 19; DB 3; Length 904;
Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCCTGCTTCCGAGC 1470
DB 642 CCCTGGCCTGCTTCCGAGC 624
```

RESULT 43
US-09-772-719B-58/c
; Sequence 58, Application US/09772719B
; Patent No. 6770438
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; Pastorekova, Silvia
; Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 465 California Street, Suite 450
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/772,719B
; FILING DATE: 30-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3A-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 904 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-772-719B-58
Query Match 1.0%; Score 19; DB 4; Length 904;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1452 CCCTGGCGCTCTCCAGC 1470
DB 642 CCTGGCGCTCTCCAGC 624
RESULT 44
US-09-270-767-11903
; Sequence 11903, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11903
; LENGTH: 988
; TYPE: DNA
; ORGANISM: Drosophila melanogaster

US-09-270-767-11903
Query Match 1.0%; Score 19; DB 4; Length 988;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 246 GCTGAGCTCTGCGCTTT 264
DB 498 GCTGAGCTCTGCGCTTT 516
RESULT 45
US-09-799-451-383
; Sequence 383, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungting
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 383
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1178)
US-09-799-451-383
Query Match 1.0%; Score 19; DB 4; Length 1742;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 148 GCCCGTCTGGAGGAGG 166
DB 891 GCCCGTCTGGAGGAGG 909
RESULT 46
US-09-252-991A-5836/c
; Sequence 5836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5836
 ; LENGTH: 2019
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-5836

Query Match 1.0%; Score 19; DB 4; Length 2019;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 CCTGCTGAGCTGGAGGCC 345
 DB 1816 CCTGCTGAGCTGGAGGCC 1798

RESULT 47
 US-09-394-142B-5
 ; Sequence 5, Application US/09394142B
 ; Patent No. 6696558
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; APPLICANT: Takayama, Shinichi
 ; TITLE OF INVENTION: No. 6696558e1 BAG Proteins and Nucleic Acid Molecules Encoding
 ; FILE REFERENCE: P-LJ 3737
 ; CURRENT APPLICATION NUMBER: US/09/394,142B
 ; PRIOR FILING DATE: 1999-09-09
 ; PRIOR APPLICATION NUMBER: 09/150,489
 ; PRIOR FILING DATE: 1998-09-09
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 2528
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2031)
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(2528)
 ; OTHER INFORMATION: n= a, c, t or g
 ; US-09-394-142B-5

Query Match 1.0%; Score 19; DB 4; Length 2528;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 631 CGGCGGCGAGAGGCC 649
 DB 207 CGGCGGCGAGAGGCC 225

RESULT 48
 US-09-394-142B-19
 ; Sequence 19, Application US/09394142B
 ; Patent No. 6696558
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; APPLICANT: Takayama, Shinichi
 ; TITLE OF INVENTION: No. 6696558e1 BAG Proteins and Nucleic Acid Molecules Encoding
 ; FILE REFERENCE: P-LJ 3737
 ; CURRENT APPLICATION NUMBER: US/09/394,142B
 ; PRIOR FILING DATE: 1999-09-09
 ; PRIOR APPLICATION NUMBER: 09/150,489
 ; PRIOR FILING DATE: 1998-09-09
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 19
 ; LENGTH: 2534
 ; TYPE: DNA

; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (307)..(2034)
 ; US-09-394-142B-19

Query Match 1.0%; Score 19; DB 4; Length 2534;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 631 CGGCGGCGAGAGGCC 649
 DB 207 CGGCGGCGAGAGGCC 225

RESULT 49
 US-09-252-991A-5773
 ; Sequence 5773, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 5773
 ; LENGTH: 2748
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-5773

Query Match 1.0%; Score 19; DB 4; Length 2748;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 CCTGCTGAGCTGGAGGCC 345
 DB 1086 CCTGCTGAGCTGGAGGCC 1104

RESULT 50
 US-08-787-739-90/c
 ; Sequence 90, Application US/08787739
 ; Patent No. 6027887
 ; GENERAL INFORMATION:
 ; APPLICANT: Zavada, Jan
 ; APPLICANT: Pastorekova, Silvia
 ; APPLICANT: Pastorek, Jaromir
 ; TITLE OF INVENTION: MN Gene and Protein
 ; NUMBER OF SEQUENCES: 96
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Leona V. Lauder
 ; STREET: 369 Pine Street, Suite 610
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/787,739
 ; FILING DATE: 24-JAN-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/485,049

;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/486,756
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/477,504
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/481,658
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/485,862
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/485,863
;; FILING DATE: 07-JUN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/487,077
;; FILING DATE: 07-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lauder, Leona L.
;; REGISTRATION NUMBER: 30,863
;; REFERENCE/DOCKET NUMBER: D-0021.4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-981-2034
;; TELEFAX: 415-981-0332
;; INFORMATION FOR SEQ ID NO: 90:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3532 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; DESCRIPTION: position 7 to position 3538 of
;; DESCRIPTION: Figure 2A-2F
;; US-08-787-739-90

Query Match 1.0%; Score 19; DB 3; Length 3532;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 CCCTGGCTGCTGCCAGC 1470
Db 3209 CCCTGGCTGCTGCCAGC 3191

Search completed: March 24, 2005, 03:14:52
Job time : 343.911 secs